

AN ANALYTIC APPROACH TO SPARSE HYPERGRAPHS: HYPERGRAPH REMOVAL

HENRY TOWNSNER

ABSTRACT. The use of tools from analysis to approach problems in graph theory has become an active area of research. Usually such methods are applied to problems involving dense graphs and hypergraphs; here we give the an extension of such methods to sparse but pseudorandom hypergraphs. We use this framework to give a proof of hypergraph removal for sub-hypergraphs of sparse random hypergraphs.

1. INTRODUCTION

In this paper we attempt to bring together two recent trends in extremal graph theory: the study of “sparse random” analogs of density theorems, and the use of methods from analysis and logic to handle complex dependencies of parameters.

To illustrate these methods, we will prove a version of the Hypergraph Removal Lemma for dense sub-hypergraphs of sparse but sufficiently pseudorandom hypergraphs. The original removal theorem was Rusza and Szemerédi’s Triangle Removal Lemma [27], which states:

Theorem 1.1. *For any every $\epsilon > 0$, there is a $\delta > 0$ such that whenever $G \subseteq \binom{V}{2}$ is a graph with at most $\delta|V|^3$ triangles, there is a set $C \subseteq G$ with $|C| \leq \epsilon|V|^2$ such that $G \setminus C$ contains no triangles at all.*

This result was later extended to graphs other than triangles [8], and ultimately to hypergraphs [10, 15, 25]. All these arguments depend heavily on the celebrated Szemerédi Regularity Lemma [29], and its generalization, the hypergraph regularity lemma [15, 26]. (Recently, Fox [9] has given a proof of graph removal without the use of the regularity lemma, which gives better bounds as a result.)

There has been a growing interest in analytic approaches to graph theory. Probably the most widely studied approach is the method of *graph limits* and *graphons* introduced by Lovász and coauthors [2, 23, 24]. Related but distinct approaches have been studied by Hrushovski, Tao, and others [1, 13, 16]. Analytic proofs of regularity and removal lemmas have been giving using all these methods [7, 30, 31, 34]. These techniques obtain a correspondence between a sequence of arbitrarily large finite graphs on the one hand, and

Date: March 2, 2013.

Partially supported by NSF grant DMS-1157580.

some sort of infinitary structure on the other. Statements about density fit naturally in these frameworks since the normalized counting measure on large finite graphs corresponds to an ordinary measure on the infinitary structure.

In this paper, we describe a similar correspondence which applies to sub-hypergraphs of sparse, pseudorandom hypergraphs. In the finite setting, the natural replacement for the normalized counting measure is the counting measure normalized by the ambient hypergraph. This introduces new complications in the infinitary world: we end up with a natural measure on sets of k -tuples which is not a genuine product measure. (This perspective on the problem was suggested to us by Hrushovski.) In place of a single measure, we end up with a family of measures, and the pseudorandomness from the finitary setting is used to ensure that this family of measures obeys certain compatibility properties.

We use this method to give an analytic proof of sparse hypergraph removal. Our proof is heavily inspired by the combinatorial proof of triangle removal in sparse graphs [20, 21]. Very recently, combinatorial proofs of sparse hypergraph removal [5, 11, 28] have also been given by several authors.

Our approach to hypergraph removal depends heavily on the use of the Gowers uniformity (semi)norms [14]. As Conlon and Gowers point out [5], such an approach cannot hope to give optimal bounds, and, relatedly, depends on a much stronger notion of pseudorandomness than strictly needed. We stick to this method both because we believe these norms are interesting in their own right, and because we believe it illustrates the analytic approach to sparse hypergraphs more clearly than an attempt to derive optimal bounds would.

In [13], Isaac Goldbring and the author proposed a general framework for handling analytic arguments of this type, which we called *approximate measure logic*. In this paper, there is no assumption that the reader is familiar with that particular framework, but we pass quickly over the logical preliminaries, and refer the reader to that paper for more detailed exposition.

We now give a brief outline of the paper. Sections 3 and 4 culminate in a proof of the dense hypergraph removal lemma, giving an outline of how we will prove sparse hypergraph removal. The reader interested primarily in the formal framework for handling sparse hypergraphs may wish to read Section 2 (which simply introduces notational conventions we use throughout the paper) and then skip to Section 5, where we actually introduce the framework.

Section 3 introduces the σ -algebras $\mathcal{B}_{V,\mathcal{I}}$, which contain those sets of tuples which can be defined (approximately) using only certain restricted sets of coordinates. (For instance, the “cut-norm” used in the theory of graph limits is closely related to the norm of the projection of a function onto the simplest non-trivial example, $\mathcal{B}_{V,1}$.) Proofs of hypergraph removal are typically divided into a *regularity* lemma and a *counting* lemma; in Section 4 we define the notion of measure having regularity—an analog of satisfying an

infinitary analog of the regularity lemma—and then prove a counting lemma. We then show that the measures corresponding to dense hypergraphs have regularity, giving a proof of the ordinary (dense) hypergraph removal lemma. This forms the outline of our approach, and we spend the rest of the paper showing that certain sparse measures also have regularity.

Section 5 finally introduces our formal framework: we define the notion of a canonical family of measures, a collection of measures having certain joint properties, and show that the ultraproduct of sufficiently random finite graphs gives us such a family. Section 6 then introduces the Gowers uniformity seminorms and begins the project of showing that, under suitable conditions, a function has positive uniformity seminorm exactly when it correlates with certain σ -algebras. Finally, in Section 7 we complete the proof of this relationship for canonical families of measures, and show that

We are grateful to Ehud Hrushovski for providing the crucial motivating idea, and to Isaac Goldbring for many helpful discussions on this topic.

2. NOTATION

Throughout this paper we use a slightly unconventional notation for tuples which is particularly conducive to our arguments. When V is a finite set, a V -tuple from G is a function $x : V \rightarrow G$. If for each $v \in V$ we have designated an element $x_v \in G$, we write x_V for the tuple $x_V(v) = x_v$. Conversely, if we have specified a V -tuple x_V , we often write x_v for $x_V(v)$. When V, W are disjoint sets, we write $x_V \cup x_W$ for the corresponding $V \cup W$ -tuple. When $I \subseteq V$ and x_V is a given V -tuple, we write x_I for the corresponding I -tuple: $x_I(i) = x_V(i)$ for $i \in I$. We write 0^V for the tuple which is constantly equal to 0. (This is the only constant tuple we will explicitly refer to.)

One of the key tools in this paper will be the use of the collection of definable sets in a model of first-order logic. We will refer to our models as $\mathfrak{M}, \mathfrak{N}$, and to the corresponding universes of these models as M, N respectively. We will refer to formal variables in the language of first-order logic with the letter w , reserving the letters x, y and so on for elements of models (for instance, when integrating over a model). We will often refer to fixed elements of a model (used as constants or parameters) with the letters a, b, c . In keeping with our tuple notation, we will often refer to finite sets of variables as w_V, w_W , etc..

Recall that when φ is a formula with free variables w_V , \mathfrak{M} is a model of first-order logic, and $x_V \in M^V$, we write $\mathfrak{M} \models \varphi(x_V)$ to indicate that the formula holds when we interpret each free variable w_v by the element x_v . A set $B \subseteq M^V$ is *definable* if $B = \{x_V \mid \mathfrak{M} \models \varphi(x_V)\}$ for some formula φ . When the model \mathfrak{M} is clear from context, we will often equate formulas with the sets they define—for instance, if B is a definable set, we will also consider B to be the formula defining this set, so by abuse of notation, $B = \{x_V \mid \mathfrak{M} \models B(x_V)\}$. If a set has multiple groups of parameters—say, $B \subseteq M^{W \cup V}$ —we will write $B(a_W)$ for the slice $\{x_V \mid a_W \cup x_V \in B\}$

corresponding to those coordinates. We say B is *definable from parameters* if $B = C(a_W)$ for some definable set C .

Similarly, when f is a simple function built from sets definable from parameters, so $f = \sum_i \alpha_i \chi_{C_i}$ where each α_i is rational and each C_i is definable from parameters, we sometimes view f as being a “rational linear combination” of formulas, and refer to the union of the parameters defining all the sets C_i as the parameters of f .

3. σ -ALGEBRAS

Models come equipped with certain natural σ -algebras.

Definition 3.1. Let \mathfrak{M} be a model and let V be a finite set of indices. We define \mathcal{B}_V^0 to be the Boolean algebra generated of subsets of M^V definable from parameters.

For $I \subseteq V$, we define $\mathcal{B}_{V,I}^0$ to be the Boolean algebra generated by subsets of M^n of the form

$$\{x_V \in M^V \mid x_I \in B\}$$

where $B \in \mathcal{B}_I^0$.

If $\mathcal{I} \subseteq \mathcal{P}(V)$ then we write $\mathcal{B}_{V,\mathcal{I}}^0$ for the Boolean algebra generated by $\bigcup_{I \in \mathcal{I}} \mathcal{B}_{V,I}^0$. When $k \leq |V|$, we define $\mathcal{B}_{V,k}^0$ to be the Boolean algebra $\mathcal{B}_{V,\{I \subseteq V \mid |I|=k\}}^0$.

For any $I \subseteq V$, we write $<I$ for the set of proper subsets of I . The *principal* algebras are those of the form $\mathcal{B}_{V,<V} = \mathcal{B}_{V,|V|-1}$.

In all cases, we drop the superscript 0 to indicate the σ -algebra generated by the algebra.

The algebras $\mathcal{B}_{V,\mathcal{I}}^0$ are generally uncountable, and so the corresponding σ -algebras $\mathcal{B}_{V,\mathcal{I}}$ are generally non-separable. It is possible to recover separability by allowing only formulas whose parameters come from an elementary submodel. (This causes some additional complications, since the slices of some set $A \subseteq M^2$ are no longer necessarily measurable; rather, the slices are measurable with respect to some slightly larger σ -algebra which depends on the choice of slice. These complications can be addressed by a small amount of additional model-theoretic work; this separable approach is used in [13,34].) These σ -algebras are closely related to the Szemerédi Regularity Lemma; for instance, in [13] it is shown that the usual regularity lemma follows almost immediately from the existence of the projection of a set onto $\mathcal{B}_{\{1,2\},1}$.

Note that, while a σ -algebra is well-defined independently of the choice of a particular measure, notions like the projection onto a σ -algebra do depend on a particular choice of measure.

The simplest interesting case of these algebras is $\mathcal{B}_{\{1,2\},1}$, which is a σ -algebra on M^2 consisting of sets of pairs which can be “defined one coordinate at a time”— $\mathcal{B}_{\{1,2\},1}$ is generated by sets of the form

$$\{(x, y) \mid \mathfrak{M} \models \varphi(x)\psi(y)\}.$$

(The first introduction of these algebras is that we know of is in [32], where Tao already notes the relationship with the Gowers uniformity norms which we will discuss in detail below. The work in this paper builds on further developments in [31, 33].)

There is some flexibility in the choice of the set \mathcal{I} ; for instance, $\mathcal{B}_{\{1,2,3\},\{\{1,2\}\}} = \mathcal{B}_{\{1,2,3\},\{\{1,2\},\{1\}\}}$ (since $\{1,2\} \in \mathcal{I}$, we may already use formulas which refer only to the coordinate 1, so adding $\{1\}$ does nothing). This leads to two canonical choices for \mathcal{I} : a minimal choice with only the sets of coordinates absolutely necessary, or a maximal choice which adds every set of coordinates allowed without changing the meaning. Depending on the situation, we want one or the other canonical form.

Lemma 3.2. *If for every $I \in \mathcal{I}$ there is an $I' \in \mathcal{I}'$ with $I \subseteq I'$ then $\mathcal{B}_{V,\mathcal{I}} \subseteq \mathcal{B}_{V,\mathcal{I}'}$.*

Proof. It suffices to show that if $I \subseteq I'$ then $\mathcal{B}_{V,I}^0 \subseteq \mathcal{B}_{V,I'}^0$. But this is easily seen from the definition, since a formula containing only the variables w_I certainly contains only the variables $w_{I'}$. \square

Corollary 3.3. *For any V, \mathcal{I} , there exist $\mathcal{I}_0, \mathcal{I}_1$ such that:*

- (1) $\mathcal{B}_{V,\mathcal{I}} = \mathcal{B}_{V,\mathcal{I}_0} = \mathcal{B}_{V,\mathcal{I}_1}$,
- (2) \mathcal{I}_0 is downwards closed: if $I \in \mathcal{I}_0$ and $J \subseteq I$ then $J \in \mathcal{I}_0$,
- (3) If $I, J \in \mathcal{I}_1$ then $J \not\subseteq I$.

Definition 3.4. Given $\mathcal{I}, \mathcal{J} \subseteq \mathcal{P}(V)$, we define $\mathcal{I} \wedge \mathcal{J}$ to consist of those sets K such that there is an $I \in \mathcal{I}$ and a $J \in \mathcal{J}$ such that $K \subseteq I \cap J$.

4. HYPERGRAPH REMOVAL

In this section, we present a proof of the ordinary hypergraph removal theorem, essentially the one given in [34], which is in turn based on the arguments in [30, 31]. We first state a necessary property on measures, and prove a lemma reminiscent of the hypergraph counting lemma.

Definition 4.1. Let ν^V be a probability measure on \mathcal{B}_V . We say ν^V has *J-regularity* for $J \subseteq V$ if:

Suppose $\mathcal{I} \subseteq \mathcal{P}(V)$ and for each $I \in \mathcal{I}$, $I \cap J \subsetneq I$. For each $I \in \mathcal{I}$, let $f_I \in L^\infty(\mathcal{B}_{V,I})$. Then if $g \in L^\infty(\mathcal{B}_{V,J})$,

$$\int (g - \mathbb{E}(g \mid \mathcal{B}_{V,\mathcal{I} \wedge J})) \prod_{I \in \mathcal{I}} f_I d\nu^V = 0.$$

When $J = V$, this is trivial.

Theorem 4.2. *Suppose ν^V has J-regularity for all $J \subseteq V$ with $|J| \leq k$, that $k < |V|$, $\mathcal{I} \subseteq \binom{V}{k}$, and for each $I \in \mathcal{I}$ we have a set $A_I \in \mathcal{B}_{V,I}$. Further, suppose there is a $\delta > 0$ such that whenever $B_I \in \mathcal{B}_{V,I}^0$ and $\nu^V(A_I \setminus B_I) < \delta$ for all $I \in \mathcal{I}$, $\bigcap_{I \in \mathcal{I}} B_I$ is non-empty. Then $\nu^V(\bigcap_{I \in \mathcal{I}} A_I) > 0$.*

Proof. We proceed by main induction on k . When $k = 1$, the claim is trivial: we must have $\nu^V(A_I) > 0$ for all I , since otherwise we could take $B_I = \emptyset$; then $\nu^V(\bigcap A_I) = \prod \nu^V(A_I) > 0$. So we assume that $k > 1$ and that whenever $B_I \in \mathcal{B}_{V,I}$ and $\nu^V(A_I \setminus B_I) < \delta$ for all I , $\bigcap_{I \in \mathcal{I}} B_I$ is non-empty. Throughout this proof, the variables I and I_0 range over \mathcal{I} .

Claim 1. For any I_0 , there is an $A'_{I_0} \in \mathcal{B}_{V,<I_0}$ such that:

- Whenever $B_I \in \mathcal{B}_{V,I}^0$ for each I , $\nu^V(A_I \setminus B_I) < \delta$ for each $I \neq I_0$, and $\nu^V(A'_{I_0} \setminus B_{I_0}) < \delta$, $\bigcap_{I \in \mathcal{I}} B_I$ is non-empty, and
- If $\nu^V(A'_{I_0} \cap \bigcap_{I \neq I_0} A_I) > 0$, $\nu^V(\bigcap_{I \in \mathcal{I}} A_I) > 0$.

Proof. Define $A'_{I_0} := \{x_{I_0} \mid \mathbb{E}(\chi_{A_{I_0}} \mid \mathcal{B}_{V,<I_0})(x_{I_0}) > 0\}$. If $\nu^V(A'_{I_0} \cap \bigcap_{I \neq I_0} A_I) > 0$ then we have

$$\int \mathbb{E}(\chi_{A_{I_0}} \mid \mathcal{B}_{V,<I_0}) \prod_{I \neq I_0} \chi_{A_I} d\nu^V > 0$$

and since ν^V has I_0 -regularity, this implies that $\nu^V(\bigcap A_I) > 0$.

Suppose that for each I , $B_I \in \mathcal{B}_{V,I}^0$ with $\nu^V(A_I \setminus B_I) < \delta$ for $I \neq I_0$ and $\nu^V(A'_{I_0} \setminus B_{I_0}) < \delta$. Since

$$\nu^V(A_{I_0} \setminus A'_{I_0}) = \int \chi_{A_{I_0}} (1 - \chi_{A'_{I_0}}) d\nu^V = \int \mathbb{E}(\chi_{A_{I_0}} \mid \mathcal{B}_{V,<I_0}) (1 - \chi_{A'_{I_0}}) d\nu^V = 0,$$

we have $\nu^V(A_{I_0} \setminus B_{I_0}) < \delta$ as well, and therefore $\bigcap_{I \in \mathcal{I}} B_I$ is non-empty. \dashv

By applying the previous claim to each $I \in \mathcal{I}$, we may assume for the rest of the proof that for each I , $A_I \in \mathcal{B}_{V,<I}$.

Fix some finite algebra $\mathcal{B} \subseteq \mathcal{B}_{V,k-1}^0$ so that for every I , $\|\chi_{A_I} - \mathbb{E}(\chi_{A_I} \mid \mathcal{B})\|_{L^2(\nu^V)} < \frac{\sqrt{\delta}}{\sqrt{2}(|\mathcal{I}|+1)}$ (such a \mathcal{B} exists because there are finitely many I and each A_I is $\mathcal{B}_{V,k-1}$ -measurable). For each I , set $A_I^* = \{a_I \mid \mathbb{E}(\chi_{A_I} \mid \mathcal{B})(a_I) > \frac{|\mathcal{I}|}{|\mathcal{I}|+1}\}$.

Claim 2. For each I , $\nu^V(A_I \setminus A_I^*) \leq \delta/2$

Proof. $A_I \setminus A_I^*$ is the set of points such that $\chi_{A_I} - \mathbb{E}(\chi_{A_I} \mid \mathcal{B})(\vec{a}) > \frac{1}{|\mathcal{I}|+1}$. By Chebyshev's inequality, the measure of this set is at most

$$(|\mathcal{I}|+1)^2 \int (\chi_{A_I} - \mathbb{E}(\chi_{A_I} \mid \mathcal{B}))^2 d\nu^V = (|\mathcal{I}|+1)^2 \|\chi_{A_I} - \mathbb{E}(\chi_{A_I} \mid \mathcal{B})\|_{L^2(\nu^V)}^2 < \frac{\delta}{2}.$$

\dashv

Claim 3. $\nu^V(\bigcap_I A_I) \geq \nu^V(\bigcap_I A_I^*) / (|\mathcal{I}| + 1)$.

Proof. For each I_0 ,

$$\begin{aligned}
\nu^V((A_{I_0}^* \setminus A_{I_0}) \cap \bigcap_{I \neq I_0} A_I^*) &= \int \chi_{A_{I_0}^*} (1 - \chi_{A_{I_0}}) \prod_{I \neq I_0} \chi_{A_I^*} d\nu^V \\
&= \int \chi_{A_{I_0}^*} (1 - \mathbb{E}(\chi_{A_{I_0}} \mid \mathcal{B})) \prod_{I \neq I_0} \chi_{A_I^*} d\nu^V \\
&\leq \frac{1}{|\mathcal{I}| + 1} \int \prod_{I \in \mathcal{I}} \chi_{A_I^*} d\nu^V \\
&= \frac{1}{|\mathcal{I}| + 1} \nu^V\left(\bigcap_{I \in \mathcal{I}} A_I^*\right)
\end{aligned}$$

But then

$$\nu^V\left(\bigcap_{I \in \mathcal{I}} A_I^* \setminus \bigcap_{I \in \mathcal{I}} A_I\right) \leq \sum_{I_0} \nu^V((A_{I_0}^* \setminus A_{I_0}) \cap \bigcap_{I \neq I_0} A_I^*) \leq \frac{|\mathcal{I}|}{|\mathcal{I}| + 1} \nu^V\left(\bigcap_{I \in |\mathcal{I}|} A_I^*\right).$$

⊥

Each A_I^* may be written in the form $\bigcup_{i \leq r_I} A_{I,i}^*$ where $A_{I,i}^* = \bigcap_{J \in \binom{I}{k-1}} A_{I,i,J}^*$ and $A_{I,i,J}^*$ is an element of $\mathcal{B}_{V,J}^0$. We may assume that if $i \neq i'$ then $A_{I,i}^* \cap A_{I,i'}^* = \emptyset$.

We have

$$\nu^V\left(\bigcap_I A_I^*\right) = \nu^V\left(\bigcup_{\vec{i} \in \prod_I [1, r_I]} \bigcap_I \bigcap_{J \in \binom{I}{k-1}} A_{i_I, J, I}^*\right).$$

For each $\vec{i} \in \prod_I [1, r_I]$, let $D_{\vec{i}} = \bigcap_I \bigcap_{J \in \binom{I}{k-1}} A_{i_I, J, I}^*$. Each $A_{I,i_I, J}^*$ is an element of $\mathcal{B}_{V,J}^0$, so we may group the components and write $D_{\vec{i}} = \bigcap_{J \in \binom{V}{k-1}} D_{\vec{i}, J}$ where $D_{\vec{i}, J} = \bigcap_{I \supset J} A_{i_I, J, I}^*$.

Suppose, for a contradiction, that $\nu^V(\bigcap_I A_I^*) = 0$. Then for every $\vec{i} \in \prod_I [1, r_I]$, $\nu^V(D_{\vec{i}}) = \nu^V(\bigcap_J D_{\vec{i}, J}) = 0$. By the inductive hypothesis, for each $\gamma > 0$, there is a collection $B_{\vec{i}, J} \in \mathcal{B}_{V,J}^0$ such that $\nu^V(D_{\vec{i}, J} \setminus B_{\vec{i}, J}) < \gamma$ and $\bigcap_J B_{\vec{i}, J} = \emptyset$. In particular, this holds with $\gamma = \frac{\delta}{2 \binom{k}{k-1} (\prod_I r_I) (\max_I r_I)}$.

For each $I, i \leq r_I, J \subset I$, define

$$B_{I,i,J}^* = A_{I,i,J}^* \cap \bigcap_{\vec{i}, i_I = i} \left[B_{\vec{i}, J} \cup \bigcup_{I' \supseteq J, I' \neq I} \overline{A_{I', i_{I'}, J}^*} \right].$$

Claim 4. $\nu^V(A_{I,i,J}^* \setminus B_{I,i,J}^*) \leq \frac{\delta}{2 \binom{k}{k-1} (\max_I r_I)}$.

Proof. Observe that if $x \in A_{I,i,J}^* \setminus B_{I,i,J}^*$ then for some \vec{i} with $i_I = i$, $x \notin B_{\vec{i}, J} \cup \bigcup_{I' \supseteq J, I' \neq I} \overline{A_{I', i_{I'}, J}^*}$. This means $x \notin B_{\vec{i}, J}$ and $x \in \bigcap_{I' \supseteq J} A_{I', i_{I'}, J}^* = D_{\vec{i}, J}$.

So

$$\nu^V(A_{I,i,J}^* \setminus B_{I,i,J}^*) \leq \sum_{\vec{i} \in \prod_I [1, r_I]} \nu^V(D_{\vec{i},J} \setminus B_{\vec{i},J}) \leq \frac{\delta}{2 \binom{k}{k-1} (\max_I r_I)}.$$

⊣

Define $B_I^* = \bigcup_{i \leq r_I} \bigcap_J B_{I,i,J}^*$.

Claim 5. $\nu^V(A_I \setminus B_I^*) < \delta$.

Proof. Since $\nu^V(A_I \setminus A_I^*) < \delta/2$, it suffices to show that $\nu^V(A_I^* \setminus B_I^*) < \delta/2$.

$$\begin{aligned} \nu^V(A_I^* \setminus \bigcup_i \bigcap_J B_{I,i,J}^*) &= \nu^V \left(\bigcup_i \bigcap_J A_{I,i,J}^* \setminus \bigcup_i \bigcap_J B_{I,i,J}^* \right) \\ &\leq \sum_{i \leq r_I} \nu^V \left(\bigcap_J A_{I,i,J}^* \setminus \bigcap_J B_{I,i,J}^* \right) \\ &\leq \sum_{i \leq r_I} \sum_J \nu^V(A_{I,i,J}^* \setminus B_{I,i,J}^*) \\ &\leq r_I \cdot \binom{k}{k-1} \cdot \frac{\delta}{2 \binom{k}{k-1} (\max_I r_I)} \\ &\leq \delta/2 \end{aligned}$$

⊣

Claim 6.

$$\bigcap_I B_I^* \subseteq \bigcup_{\vec{i}} \bigcap_J B_{\vec{i},J}.$$

Proof. Suppose $x \in \bigcap_I B_I^* = \bigcap_I \bigcup_{i \leq r_I} \bigcap_J B_{I,i,J}^*$. Then for each I , there is an $i_I \leq r_I$ such that $x \in \bigcap_J B_{I,i_I,J}^*$. Since $B_{I,i_I,J}^* \subseteq A_{I,i_I,J}^*$, for each I and $J \subset I$, $x \in A_{I,i_I,J}^*$.

For any J , let $I \supset J$. Then

$$x \in B_{I,i_I,J}^* = A_{I,i_I,J}^* \cap \bigcap_{\vec{i}', i'_I = i_I} \left[B_{\vec{i}',J} \cup \bigcup_{I' \supseteq J, I' \neq I} \overline{A_{I',i_{I'},J}^*} \right].$$

In particular, $x \in \left[B_{\vec{i},J} \cup \bigcup_{I' \supseteq J, I' \neq I} \overline{A_{I',i_{I'},J}^*} \right]$ for the particular \vec{i} we have chosen. Since $x \in A_{I,i_{I'},J}^*$ for each $I' \supset J$, it must be that $x \in B_{\vec{i},J}$. This holds for any J , so $x \in \bigcap_J B_{\vec{i},J}$. ⊣

From our assumption, $\bigcap_I B_I^*$ is non-empty, and therefore there is some \vec{i} such that $\bigcap_J B_{\vec{i},J}$. But this leads to a contradiction, so it must be that $\nu^V(\bigcap_I A_I^*) > 0$, and therefore, as we have shown, $\nu^V(\bigcap_{I \in \mathcal{I}} A_I) \geq \frac{1}{|\mathcal{I}|+1} \nu^V(\bigcap_{I \in \mathcal{I}} A_I^*) > 0$. □

In order to prove the hypergraph removal theorem, we would then hope to argue as follows: the failure of hypergraph removal implies the existence of a family of counterexamples of unbounded size. We could then take the ultraproduct of these counterexamples to obtain an infinite model in which $\nu^V(\bigcap_{I \in \binom{V}{k}} A_I) = 0$ for a family of sets A_I corresponding to the graph we are trying to remove. By the previous theorem, we would have an arbitrarily small family of definable sets B_I , and we would then argue that that these sets correspond to sets in the finite models whose removal causes the removal of all copies of the hypergraph. The only remaining difficulty in this argument is showing that the measure we obtain has J -regularity for all $J \subseteq V$.

In the remainder of this section, we carry out the proof for the dense case.

Definition 4.3. Let $\varphi(w_V, w_W)$ be a formula with the displayed free variables. Then there is a corresponding function $A_\varphi : M^W \rightarrow \mathcal{B}_V$ given by $A_\varphi(x_W) = \{x_V \mid \mathfrak{M} \models \varphi(x_V, x_W)\}$. Let ν be a probability measure on \mathcal{B}_V . For any φ , define $\nu_\varphi : M^W \rightarrow [0, 1]$ by $\nu_\varphi(x_W) = \nu(A_\varphi(x_W))$. We say ν is an *definable Keisler probability measure* if for every formula φ with parameters, ν_φ is continuous with respect to the topology generated by \mathcal{B}_W^0 .

Lemma 4.4. Suppose that for each $J \subseteq V$, ν^J is a definable Keisler probability measure on \mathcal{B}_J such that for any $L^\infty(\nu^V)$ function f , $\int f d\nu^V = \iint f d\nu^J d\nu^{V \setminus J}$. Then ν^V has J -regularity for every $J \subseteq V$.

Proof. We have

$$\begin{aligned} & \int (g - \mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I} \wedge J})) \prod_I f_I d\nu^V \\ &= \int (g - \mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I} \wedge J})) \prod_I f_I d\nu^J d\nu^{V \setminus J}. \end{aligned}$$

For each $a_{V \setminus J}$, the function $\prod_I f_I(a_{I \setminus J}, x_{I \cap J})$ is measurable with respect to $\mathcal{B}_{V, < J}$, so we have

$$\int (g - \mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I} \wedge J})) \prod_I f_I d\nu^J = 0.$$

Since this holds for every $a_{V \setminus J}$, the claim follows by integrating over all choices of $a_{V \setminus J}$. \square

Definition 4.5. Let K, A be k -uniform hypergraphs on vertex sets $V(K), V(A)$ respectively. $\pi : V(K) \rightarrow V(A)$ is a *homomorphism* if whenever $e \in K$, $\pi''e \in A$. (That is, π maps edges to edges.) $\text{hom}(K, A)$ is the number of distinct homomorphisms from K to A . If K, A are k -uniform hypergraphs, we write

$$d(K, A) = \frac{\text{hom}(K, A)}{|V(A)|^{|V(K)|}}.$$

Theorem 4.6. *For every k -uniform hypergraph K and constant $\epsilon > 0$, there is a δ such that whenever A is a finite k -uniform hypergraph with $d(K, A) < \delta$, there is a subset L of A with $|C| \leq \epsilon \binom{|V(A)|}{k}$ such that $\text{hom}(K, A \setminus C) = 0$.*

Proof. Suppose not. Let K, ϵ be a counterexample, and since there is no such δ , for each n we may choose a k -uniform hypergraph A^n with $d(K, A^n) < 1/n$ such that there is no such subset L of A^n . Clearly $|V(A^n)| \rightarrow \infty$. We view each A^n as a model, with $M^n = V(A^n)$ the set of points, A^n a k -ary relation $V(A^n)$, and predicates making the normalized counting measure ν_n^J on $V(A^n)^J$ definable for each $J \subseteq V(K)$. In particular, this means the counting measure is a *uniformly* definable Keisler probability measure.

Let $V = V(K)$. For each $I \in K$, let $A_I^n = \{x_V \mid x_I \in A^n\}$. Note that $\text{hom}(K, A^n)$ consists exactly of the elements of

$$\bigcap_{I \in K} A_I^n,$$

and therefore $d(K, A^n) = \nu_n^V(\bigcap_{I \in K} A_I^n)$. In particular, we have $\nu_n^V(\bigcap_{I \in K} A_I^n) \rightarrow 0$.

Now take an ultraproduct of the models (M^n, A^n, \dots) to obtain $\mathfrak{M} = (M, A, \dots)$. (See [12] for the construction and, in particular, the demonstration that the measures defined by ν^J , the ultraproduct of the ν_n^J , extend to probability measures on \mathcal{B}_J .) By [17, 18] (or see the next section), the decomposition in the statement of Lemma 4.4 holds in \mathfrak{M} , and therefore ν^V has I -regularity for all $I \subseteq V$. We have $\nu^V(\bigcap_{I \in K} A_I) = 0$, and therefore by the previous theorem, there are $B_I \in \mathcal{B}_{V, I}^0$ such that $\nu^V(E \setminus B_I) < \frac{\epsilon}{|K|}$ and $\bigcap_{I \in K} B_I = \emptyset$. Let $C = \bigcup_I (A_I \setminus B_I)$, so $\nu^{[1, k]}(C) < \epsilon$. C is definable from parameters in M , and therefore

$$\bigcap_{I \in K} (A_I \setminus C) = \emptyset$$

is a formula, which is therefore satisfied by the corresponding set in almost every (M^n, A^n, \dots) . Let C^n be the set defined in the model (M^n, A^n, \dots) by the formula defining C . Then there is some sufficiently large n such that $\nu_n^V(C^n) < \epsilon$ but $\bigcap_{I \in K} (A_I^n \setminus C^n) = \emptyset$, contradicting the assumption. \square

Our goal is to obtain the same result when A is not a dense hypergraph, but rather a dense subset of a sparse random graph. The main idea is that we will replace ν^V with a measure concentrating on the sparse random graph; however this will not satisfy the easy Fubini decomposition we used for the dense case, so we will need to use the randomness—plus a large amount of additional machinery—to prove that the resulting measures nonetheless have regularity.

5. FAMILIES OF MEASURES

To motivate our construction, we first consider the situation in large finite graphs. Suppose we have a large finite set of vertices G and a sparse random

graph Γ on G . There are two natural measures we might consider on subsets of G^2 : the usual normalized counting measure

$$\lambda(S) = \frac{|S|}{|G|^2}$$

and the counting measure normalized by Γ :

$$\lambda'(S) = \frac{|S \cap \Gamma|}{|\Gamma|}.$$

When we consider subsets of G^3 , we have even more choices; we could normalize with respect to all possible triangles

$$\lambda_0(S) = \frac{|S|}{|G|^3},$$

or only those triangles entirely in Γ

$$\lambda_1(S) = \frac{|\{(x, y, z) \in S \mid (x, y) \in \Gamma, (x, z) \in \Gamma, (y, z) \in \Gamma\}|}{|\{(x, y, z) \mid (x, y) \in \Gamma, (x, z) \in \Gamma, (y, z) \in \Gamma\}|},$$

or only those triangles where certain specified edges belong to Γ :

$$\lambda_2(S) = \frac{|\{(x, y, z) \in S \mid (x, y) \in \Gamma, (x, z) \in \Gamma\}|}{|\{(x, y, z) \mid (x, y) \in \Gamma, (x, z) \in \Gamma\}|}.$$

Indeed, further consideration suggests that we have multiple choices for measures even on subsets of G : in addition to the normalized counting measure, we could fix any element $x \in G$ and define

$$\lambda_x(S) = \frac{|\{y \in S \mid (x, y) \in \Gamma\}|}{|\{y \mid (x, y) \in \Gamma\}|}.$$

When Γ is a k -uniform hypergraph, we have yet more possibilities.

We therefore introduce a general notation for referring to all such measures. We first describe this notation in the setting of a large finite graph, but we will primarily use it in the infinitary setting. We assume that a value for k and a k -uniform hypergraph Γ have been fixed. We let V be a finite set of indices, and we describe a family of measures on G^V : let W be a set disjoint from V , let $x_W \in G^W$, and let E be a k -uniform hypergraph on $V \cup W$. Define

$$\Gamma_{E, x_W}^V = \{x_V \in G^V \mid \forall e \in E \ x_e \in \Gamma\}.$$

Note the significance of our notation for tuples here: x_e is a k -tuple which may consist both of elements from the fixed set x_W and from x_V . E specifies which sets of k indices from $V \cup W$ are required to belong to Γ . For instance, in the case where $k = 2$, so Γ is a graph, $\Gamma_{\{(1,2)\}, \emptyset}^{\{1,2\}} = \Gamma$, while $\Gamma_{\emptyset, \emptyset}^{\{1,2\}} = G^2$.

We then define

$$\mu_{E, x_W}^V(S) = \begin{cases} \frac{|S \cap \Gamma_{E, x_W}^V|}{|\Gamma_{E, x_W}^V|} & \text{if } |\Gamma_{E, x_W}^V| > 0 \\ 0 & \text{if } |\Gamma_{E, x_W}^V| = 0 \end{cases}.$$

For instance, in the measures above, $\lambda = \mu_{\emptyset, \emptyset}^{\{1,2\}}$, $\lambda' = \mu_{\{(1,2)\}, \emptyset}^{\{1,2\}}$, $\lambda_0 = \mu_{\emptyset, \emptyset}^{\{1,2,3\}}$, $\lambda_1 = \mu_{\{(1,2), (1,3), (2,3)\}, \emptyset}^{\{1,2,3\}}$, $\lambda_2 = \mu_{\{(1,2), (1,3)\}, \emptyset}^{\{1,2,3\}}$, and $\lambda_x = \mu_{\{(1,2)\}, x}^{\{1\}}$. When W and x_W are clear from context, we just write μ_E^V for μ_{E, x_W}^V , and call x_W the *back-ground parameters* of μ_E^V .

When integrating over μ_{E, x_W}^V , we always assume the variable being integrated is x_V .

A key feature of this notation is that it makes it easy to specify the Fubini-type properties that we would like these measures to satisfy. If $V = V_0 \cup V_1$ where $V_0 \cap V_1 = \emptyset$ and E' is the restriction of E to vertices from $V_0 \cup W$, we intend to have

$$\int \cdot d\mu_{E, x_W}^V = \iint \cdot d\mu_{E, x_{V_0} \cup x_W}^{V_1} d\mu_{E', x_W}^{V_0}.$$

To avoid having to endlessly specify the restriction of E to the appropriate vertices, we will generally allow E to have extra edges not included in the vertex set V ; for instance, we will not distinguish between $\mu_{E', x_W}^{V_0}$ and $\mu_{E, x_W}^{V_0}$, and will usually write

$$\int \cdot d\mu_{E, x_W}^V(x_V) = \iint \cdot d\mu_{E, x_{V_0} \cup x_W}^{V_1} d\mu_{E, x_W}^{V_0},$$

even though E is not a subset of $\binom{V_0 \cup W}{k}$.

In our infinitary setting, we no longer have the underlying counting measures to refer to, so we will have to define formally the properties we want a family of measures to have.

We will use the meta-variable μ for a *family of probability measures*—technically, a function from appropriate finite sets to probability measures, so when μ is a family of probability measures, μ_{E, x_W}^V is an actual probability measure for suitable values of V , E , x_W .

Definition 5.1. Let \mathfrak{M} be a model. A *weakly canonical family of probability measures of degree k and size d* , μ , consists of, for any sets V, W with $V \cap W = \emptyset$, any k -uniform hypergraph E on $V \cup W$ with $|E| \leq d$, and any $x_W \in M^W$, a probability measure μ_{E, x_W}^V on \mathcal{B}_V such that:

- (1) For μ_E^W -almost-every x_W , μ_{E, x_W}^V is a definable Keisler probability measure,
- (2) If no edges in E contain both w and an element of V then $\mu_{E, x_W \cup \{w\}}^V = \mu_{E, x_W}^V$,
- (3) If $\pi : V_0 \cup W_0 \rightarrow V_1 \cup W_1$ is a bijection mapping V_0 to V_1 and W_0 to W_1 and $\pi(E_0) = E_1$ then $\mu_{E_0, x_{W_0}}^{V_0} = \mu_{E_1, x_{\pi^{-1}(W_1)}}^{V_1}$.

We say μ is a *canonical family of probability measures* if additionally

When $V = V_0 \cup V_1$, $V_0 \cap V_1 = \emptyset$, for μ_E^W -almost every x_W these measures satisfy the Fubini properties

$$\int \cdot d\mu_{E, x_W}^V = \iint \cdot d\mu_{E, x_{V_0} \cup x_W}^{V_1} d\mu_{E, x_W}^{V_0}.$$

Weak canonicity merely enforces a certain amount of uniformity on these measures—the parameter x_w only matters if there is an edge in E connecting w to V , and the measures don't depend on the particular choice of indices used. The Fubini condition is non-trivial, and it is ensuring this property that requires us to work only with sufficiently random sparse graphs.

We wish to work in models which have two additional features: first, the model actually includes formulas defining the all of the measures in the family μ . Second, the model contains extra function symbols **max** which pick out values maximizing certain integrals. (The construction of such languages has appeared a few times (see [16, 35]; a general theory of constructions of this kind is given in [13]).)

Definition 5.2. Let \mathcal{L} be a language of first-order logic containing a k -ary relation symbol γ , and let d be given. $\mathcal{L}^{\gamma,d}$ is the smallest language containing \mathcal{L} such that:

- Whenever $\varphi(w_V, w_W, w_P)$ is a formula with the displayed free variables, W is a set disjoint from V , E is a k -uniform hypergraph on $V \cup W$ with $|E| \leq d$, and $q \in [0, 1]$ is rational, there are formulas

$$m_{E,w_W}^V \leq q \cdot \varphi$$

and

$$m_{E,w_W}^V < q \cdot \varphi$$

with free variables w_W, w_P , and

- Whenever E is a k -uniform hypergraph with $\leq d$ edges on a vertex set V , $V = V_0 \cup V_1$ is a partition of V , W and P are finite sets with V, W, P pairwise disjoint, f is a rational linear combination of formulas with free variables w_W, w_V , and $\varphi(w_W, w_P, w_V)$ is a formula with the displayed free variables, for each $p \in P$ there is a function symbol $\mathbf{max}_p^{E,V_0,f,\varphi}(w_W, w_{V_0})$.

Note that the formulas $m_{E,w_W}^V \leq q \cdot \varphi$ and $m_{E,w_W}^V < q \cdot \varphi$ bind the variables w_V . We will “abbreviate” these formulas as $m_{E,w_W}^V(\varphi) \leq q$ and $m_{E,w_W}^V(\varphi) < q$ respectively. We will abbreviate $\neg m_{E,w_W}^V(\varphi) \leq q$ by $m_{E,w_W}^V(\varphi) > q$ and $\neg m_{E,w_W}^V(\varphi) < q$ by $m_{E,w_W}^V(\varphi) \geq q$. We view $\{\mathbf{max}_p^{E,V_0,f,\varphi}(w_W, w_P, w_{V_0})\}_{p \in P}$ as a tuple $\mathbf{max}_P^{E,V_0,f,\varphi}(w_W, w_{V_0})$ of function symbols.

Definition 5.3. If \mathfrak{M} is a finite model of \mathcal{L} and $\Gamma = \gamma^{\mathfrak{M}}$ is the interpretation of γ in this model, we extend \mathfrak{M} to a model $\mathfrak{M}^{\Gamma,d}$ of $\mathcal{L}^{\gamma,d}$ by interpreting

$$\mathfrak{M}^{\Gamma,d} \models m_{E,a_W}^V(B) \leq q$$

to hold iff

$$\mu_{E,a_W}^V(B) \leq q$$

whenever B is definable from parameters, and similarly for $m_{E,a_W}^V(B) < q$.

Suppose we have interpreted the formula φ and all the formulas defining the simple function f . Let B be the set defined by φ . For each $a_W \in$

$M^W, x_{V_0} \in M^{V_0}$, we choose $(\mathbf{max}_P^{E, V_0, f, \varphi}(a_W, x_{V_0}))^{\mathfrak{M}^{\Gamma, d}}$ to be some tuple b_P maximizing $\left| \int f \chi_{B(a_W, x_{V_0}, b_P)} d\mu_{E, x_{V_0}}^{V_1} \right|$.

Note that we consistently use m to refer to the formula of first-order logic describing a measure, and μ to the actual measure corresponding to m . Also, note that in the interpretation of $\mathbf{max}_P^{E, V_0, f, \varphi}(a_W, x_{V_0})$, B depends on a_W, x_V , and b_P , while f depends on only a_W and x_V .

We need to ensure we work with sufficiently random sparse hypergraphs:

Definition 5.4. Let Γ be a k -uniform hypergraph on a set of n vertices and let V, W be disjoint sets and $E \subseteq \binom{V \cup W}{k}$. We write $\mathcal{E}_{V, W, E}^\delta \subseteq \Gamma_W^E$ for the set of tuples $a_W \in \Gamma_W^E$ such that there is some partition $V = V_0 \cup V_1$ and some set $A \subseteq \Gamma_{E, a_W}^{V_0}$ such that

$$\left| \mu_{E, a_W}^{V_0}(A) - \int \mu_{E, x_{V_0} \cup a_W}^{V_0}(A) d\mu_{E, a_W}^{V_0} \right| \geq \delta.$$

Let Γ be a k -uniform hypergraph on a set of n vertices. We say Γ is δ, d -suitably random if whenever V, W are disjoint sets with $|V \cup W| \leq kd$, $|E| \leq d$, $\mu_E^W(\mathcal{E}_{V, W, E}^\delta) < \delta$.

This definition says that “most” $x_{V_0} \in \Gamma_{E, a_W}^{V_0}$ have roughly the correct number of extensions to elements of Γ_{E, a_W}^V . This can be seen as a hypergraph generalization of the notion of uniformity used in, for instance, [19] to prove versions of the regularity lemma in sparse random graphs.

Let \mathcal{L} be the language consisting of two k -ary relation symbols, γ and α .

Theorem 5.5. Let $\epsilon > 0$. Suppose that for each n , Γ_n is a δ_n, d -suitably random k -uniform hypergraph where $\delta_n \rightarrow 0$, and let $A_n \subseteq \Gamma_n$ be given with $|A_n| \geq \epsilon |\Gamma_n|$. Then each $\mathfrak{M}_n = (\Gamma_n, A_n)$ is a model of \mathcal{L} . Let \mathcal{U} be an ultrafilter on \mathbb{N} and let \mathfrak{M} be the ultraproduct of the models $\mathfrak{M}_n^{\Gamma_n, d}$. Then \mathfrak{M} is a model of $\mathcal{L}^{\gamma, d}$ such that:

- (1) $\mathfrak{M} \models \sigma$ iff for \mathcal{U} -almost-every n , $\mathfrak{M}_n^{\Gamma_n, d} \models \sigma$.
- (2) There is a canonical family of probability measures of degree k and size d , μ_{E, x_W}^V on the σ -algebra generated by the definable subsets of M^V such that whenever B is definable from parameters,

$$\mu_{E, a_W}^V(B) = \inf\{q \in \mathbb{Q}^{>0} \mid \mathfrak{M} \models m_{E, a_W}^V(B) < q\}.$$

- (3) $\mu_{\{[1, k]\}}^{[1, k]}(A) \geq \epsilon$.
- (4) Whenever E is a k -uniform hypergraph with $\leq d$ edges on a vertex set V , $V = V_0 \cup V_1$ is a partition of V , W and P are finite sets with V, W, P pairwise disjoint, f is a rational linear combination of formulas with free variables w_W, w_V , and $\varphi(w_W, w_P, w_V)$ is a formula with the displayed free variables, for every $a_W \in M^W, b_P \in$

$$M^P, x_{V_0} \in M^{V_0},$$

$$\left| \int f \chi_{B(a_W, x_{V_0}, \max_P^{E, V_0, f, \varphi}(a_W, x_{V_0}))} d\mu_{E, x_{V_0}}^{V_1} \right| \geq \left| \int f \chi_{B(a_W, x_{V_0}, b_P)} d\mu_{E, x_{V_0}}^{V_1} \right|.$$

Proof. (1) The first part is the standard Łoś Theorem for ultraproducts.

- (2) That the measure μ_{E, a_W}^V defined in this way extends to a genuine measure on \mathcal{B}_V is the standard Loeb measure construction. Weak canonicity holds in each finite model, and therefore there are formulas holding in each finite model specifying that weak canonicity holds. By the Łoś Theorem, these formulas hold in \mathfrak{M} , and therefore the measures μ_{E, a_W}^V are weakly canonical.

Note that the formulas satisfied by m_{E, a_W}^V in \mathfrak{M} and the actual measure μ_{E, a_W}^V *almost* line up: when B is definable from parameters, if $\mu_{E, a_W}^V(B) < q$ then $\mathfrak{M} \models m_{E, a_W}^V(B) < q$, but if $\mathfrak{M} \models m_{E, a_W}^V(B) < q$ then we can only be sure that $\mu_{E, a_W}^V(B) \leq q$.

To see that the measures μ_{E, a_W}^V are actually canonical, it suffices to show that for each $B \in \mathcal{B}_V^0$ and μ_E^W -almost every $x_W \in M^W$,

$$\mu_{E, x_W}^V(B(x_W)) = \int \mu_{E, x_{V_0} \cup x_W}^{V_1}(B(x_W)) d\mu_{E, x_W}^{V_0}.$$

Suppose not; then for some set B definable from parameters, there is a set of x_W of positive measure such that this equality fails. It follows that for some rational $\delta > 0$ there is a set X_0 of x_W of positive measure such that

$$\left| \mu_{E, x_W}^V(B(x_W)) - \int \mu_{E, x_{V_0} \cup x_W}^{V_1}(B(x_W)) d\mu_{E, x_W}^{V_0} \right| > \delta.$$

We need to represent the integral in this definition closely enough by a formula to let us define a set of points where this violation occurs. Consider the function $f_{x_W}(x_{V_0}) = \mu_{E, x_{V_0} \cup x_W}^{V_1}(B(x_W))$. We have $0 \leq f_{x_W}(x_{V_0}) \leq 1$.

(Roughly speaking, the problem is that integrals are not directly definable in our language, and there are “different ways” a function could have a given integral—say, by having a small number of points where the value is large, or a larger number of points where the value is smaller. However we will show that there must be a set of positive measure where the functions f_{x_W} not only all have nearly the same integral, but all these integrals can be finitely approximated using the same level sets. This will allow us to write down a formula defining a set of points of positive measure, and with the property that every point satisfying this formula belongs to X_0 .)

We may partition the interval $[0, 1]$ into finitely many intervals $I_i = [\delta_i, \delta_{i+1})$ of size $< \delta/8$ and with rational endpoints. Let us set

$\Pi_i(x_W) = \{x_{V_0} \mid f_{x_W}(x_{V_0}) \in I_i\}$ and $\pi_i(x_W) = \mu_{E,x_W}^{V_0}(\Pi_i(x_W))$, so when $x_W \in X_0$, $\sum_i \delta_i \pi_i(x_W) \leq \int f_{x_W} d\mu_{E,x_W}^{V_0} < \sum_i \delta_i \pi_i(x_W) + \delta/8$.

We choose $X_1 \subseteq X_0$ of positive measure and, for each i , an interval $J_i = (\eta_j, \eta'_j)$ with rational end points such that $\pi_i(x_W) \in J_i$ for each $x_W \in X_1$ and

$$\sum_i \delta_{i+1} \eta'_i < \sum_i \delta_i \eta_i + \delta/4.$$

Choose a rational $\sigma > 0$ very small, and let

$$\begin{aligned} \Pi'_i(x_W) = \{x_{V_0} \mid \mathfrak{M} \models m_{E,x_{V_0} \cup x_W}^{V_1}(B(x_W)) < \delta_{i+1} \\ \wedge m_{E,x_{V_0} \cup x_W}^{V_1}(B(x_W)) > \delta_i - \sigma\}. \end{aligned}$$

Then $\Pi_i(x_W) \subseteq \Pi'_i(x_W)$ and $\Pi'_i(x_W)$ is definable. Let $\pi'_i(x_W) = \mu_{E,x_W}^{V_0}(\Pi'_i(x_W))$. By choosing σ small enough, we may find a set $X_2 \subseteq X_1$ of positive measure so that for $x_W \in X_2$, each $\pi'_i(x_W) \in J_i$ as well.

Now we may consider the set Θ of x_W such that

$$\forall i \left(\mathfrak{M} \models m_{E,x_W}^{V_0}(\Pi'_i(x_W)) < \eta'_i \wedge m_{E,x_W}^{V_0}(\Pi'_i(x_W)) > \eta_i \right).$$

Note that Θ is definable from parameters and $X_2 \subseteq \Theta$.

Consider any $x_W \in \Theta$, not necessarily in X_2 . Since each $\mu_{E,x_W}^{V_0}(\Pi'_i(x_W)) \leq \eta'_i$,

$$\int \mu_{E,x_{V_0} \cup x_W}^{V_1}(B(x_W)) d\mu_{E,x_W}^{V_0} \leq \sum_i \delta_{i+1} \eta'_i < \sum_i \delta_i \eta_i + \delta/4.$$

On the other hand, since each $\mu_{E,x_W}^{V_0}(\Pi'_i(x_W)) \geq \eta_i$,

$$\int \mu_{E,x_{V_0} \cup x_W}^{V_1}(B(x_W)) d\mu_{E,x_W}^{V_0} \geq \sum_i (\delta_i - \sigma) \eta_i > \sum_i \delta_i \eta_i - \delta/4$$

(since we chose σ small enough).

So when $x_W \in \Theta$, we have

$$\sum_i (\delta_i - \sigma) \eta_i - \delta/4 < \int \mu_{E,x_{V_0} \cup x_W}^{V_1}(B) d\mu_{E,x_W}^{V_0} < \sum_i \delta_i \eta_i + \delta/4.$$

Therefore when $x_W \in X_2 \subseteq X_0 \cap \Theta$, we must have either $\mu_{E,x_W}^V(B(x_W)) < \sum_i \delta_i \eta_i - \delta/2$ or $\mu_{E,x_W}^V(B(x_W)) > \sum_i \delta_i \eta_i + \delta/2$, and therefore

$$\begin{aligned} \mathfrak{M} \models & \left(m_{E,x_W}^V(B(x_W)) < \sum_i \delta_i \eta_i - \delta/2 \right) \\ & \vee \left(m_{E,x_W}^V(B(x_W)) > \sum_i \delta_i \eta_i + \delta/2 \right). \end{aligned}$$

Let ψ be the conjunction of this formula with the formula defining Θ . Then we have $\mathfrak{M} \models \psi(x_W)$ whenever $x_W \in X_2$, and therefore $\mathfrak{M} \models m_E^W(\psi) > \zeta$ for some $\zeta > 0$. We also have that whenever $\mathfrak{M} \models \psi(x_W)$, it is actually true that $\left| \mu_{E,x_W}^{V_0}(B(x_W)) - \int \mu_{E,x_{V_0} \cup x_W}^{V_1}(B(x_W)) d\mu_{E,x_W}^{V_0} \right| > \delta$.

Since the formula $m_E^W(\psi) > \zeta$ holds in the ultraproduct, it also holds in infinitely many finite models. But any finite model where this holds fails to be ζ, d -suitably random. This contradicts the assumption that the finite models are δ_n, d -suitably random with $\delta_n \rightarrow 0$.

- (3) The third requirement follows immediately the Loś Theorem: the formula $m_{\{[1,k]\}}^{[1,k]}(A) \geq \epsilon$ holds in every finite model, and therefore in \mathfrak{M} as well, and therefore $\mu_{\{[1,k]\}}^{[1,k]}(A) \geq \epsilon$.
- (4) Fortunately, the integral in this statement does not cause as much difficulty, since we do not need to deal with it uniformly in parameters. Let $f = \sum \alpha_i \chi_{C_i}$. Whenever $\left| \int f \chi_{B(a_W, x_{V_0}, b_P)} d\mu_{E,x_{V_0}}^{V_1} \right| > \epsilon$ for some ϵ , there is a formula holding of the parameters a_W, x_{V_0}, b_P which is a conjunction of components of the form

$$m_{E,x_{V_0}}^{V_1}(C_i(a_W, x_{V_0}) \wedge B(a_W, b_P, x_{V_0})) < q$$

or negations of such components, and which implies that the integral is $\geq \epsilon$. But then this formula holds in \mathcal{U} -almost every finite model, which means that we must have $\left| \int f \chi_{B(a_W, x_{V_0}, \max_P^{E,V_0,f,\varphi}(a_W, x_{V_0}))} d\mu_{E,x_{V_0}}^{V_1} \right| \geq \epsilon$ in \mathcal{U} -almost every finite model (where a_P , etc., refer to the corresponding parameters in those finite models). But then this formula also holds in \mathfrak{M} , so $\left| \int f \chi_{B(a_W, x_{V_0}, \max_P^{E,V_0,f,\varphi}(a_W, x_{V_0}))} d\mu_{E,x_{V_0}}^{V_1} \right| \geq \epsilon$ in \mathfrak{M} . Since this holds for every $\epsilon < \left| \int f \chi_{B(a_W, x_{V_0}, b_P)} d\mu_{E,x_{V_0}}^{V_1} \right|$, it follows that

$$\left| \int f \chi_{B(a_W, x_{V_0}, \max_P^{E,V_0,f,\varphi}(a_W, x_{V_0}))} d\mu_{E,x_{V_0}}^{V_1} \right| \geq \left| \int f \chi_{B(a_W, x_{V_0}, b_P)} d\mu_{E,x_{V_0}}^{V_1} \right|.$$

□

6. UNIFORMITY SEMINORMS

6.1. Seminorms for Principal Algebras. In this section we define a family of seminorms, the Gowers uniformity seminorms, corresponding to the σ -algebras defined above. Fix disjoint sets V, P and a k -uniform hypergraph $E \subseteq \binom{V \cup P}{k}$; let $m = |E \cap \binom{P}{k}|$ and let μ be a canonical family of measures of degree k and size $\sum_{I \in E} 2^{|I \cap V|}$. Let a_P be such that the measure μ_{E,a_P}^V , and the measures we generate from it below, satisfy the appropriate Fubini properties. (We will only work with a finite family of measures, so the set of

such a_P has μ_E^P -measure 1.) To avoid repeating the background parameters a_P over and over, we will write μ_E^V as an abbreviation for μ_{E,a_P}^V .

Definition 6.1. For each $I \subseteq V$, we define $\mu_E^{V+I} = \mu_{E^{V+I}}^{(V \setminus I) \cup (\{0,1\} \times I)}$ where E^{V+I} is given as follows: for each $J \in E$ and each $\omega : J \cap I \rightarrow \{0,1\}$, there is an edge $J^\omega = (J \setminus I) \cup \{(i, \omega(i)) \mid i \in J \cap I\}$.

μ_E^{V+I} is the measure obtained by making a second copy of the indices in I . For instance, if $\mu_E^V = \mu_{\{(1,2)\}}^{\{1,2\}}$, the measure whose underlying graph has two points connected by an edge, μ_E^{V+V} is a measure on 4-tuples, where the four coordinates are $\{(1,0), (1,1), (2,0), (2,1)\}$ and there is an edge between each pair $(1,b), (2,b')$.

Note that $\mu_E^{V+\emptyset} = \mu_E^V$. For $i \in I$, $b \in \{0,1\}$, we write x_i^b in place of $x_{(i,b)}$; for instance, we write

$$\int f(x_{V \setminus I}, x_I^0, x_I^1) d\mu_E^{V+I}$$

where the variables being integrated over are exactly the ones displayed. If $\omega : I \rightarrow \{0,1\}$, we write x_I^ω for the tuple $x_I^\omega(i) = x_i^{\omega(i)}$.

Note that we chose the degree of our measure to be $\sum_{I \in E} 2^{|I \cap V|}$ because this is precisely the size needed to ensure Fubini properties for μ_E^{V+V} .

Definition 6.2. Let $f : M^V \rightarrow \mathbb{R}$ be an $L^\infty(\mu_E^V)$, \mathcal{B}_V -measurable function with $|V| = n$. Define $\| \cdot \|_{U_\infty^V(\mu_E^V)}$ by:

$$\|f\|_{U_\infty^V(\mu_E^V)} = \left(\int \prod_{\omega \in \{0,1\}^V} f(x_V^\omega) d\mu_E^{V+V} \right)^{2^{-n}}.$$

Whenever we refer to the norm $\|f\|_{U_\infty^V(\mu_E^V)}$, we assume that f is $L^\infty(\mu_E^V)$ and \mathcal{B}_V -measurable.

We have to check that these are well-defined. We actually prove the following stronger lemma, which will be useful later.

Lemma 6.3. *If f is an $L^\infty(\mu_E^V)$ function and B is $\mathcal{B}_{V,I}$ -measurable for some $I \subseteq V$ then*

$$0 \leq \int \prod_{\omega \in \{0,1\}^V} f(x_V^\omega) \chi_B(x_V^\omega) d\mu_E^{V+V} \leq \int \prod_{\omega \in \{0,1\}^V} f(x_V^\omega) d\mu_E^{V+V}.$$

Proof. It suffices to show the claim in the case when $|I| = |V| - 1$. Since $f = f\chi_B + f\chi_{\overline{B}}$, we have

$$\int \prod_{\omega \in \{0,1\}^V} f(x_V^\omega) d\mu_E^{V+V} = \int \prod_{\omega \in \{0,1\}^V} [(f\chi_B)(x_V^\omega) + (f\chi_{\overline{B}})(x_V^\omega)] d\mu_E^{V+V}.$$

Expanding the product gives a sum of 2^{2^n} terms of the form

$$\int \prod_{\omega \in \{0,1\}^V} (f\chi_{S_\omega})(x_V^\omega) d\mu_E^{V+V}$$

where each S_ω is either χ_B or $\chi_{\overline{B}}$. We will show that each of these terms is non-negative. Since $\int \prod_{\omega \in \{0,1\}^V} f(x_V^\omega) \chi_B(x_V^\omega) d\mu_E^{V+V}$ is one of these terms, both inequalities follow.

Note that $\chi_{S_\omega}(x_V^\omega)$ depends only on x_I^ω . In particular, if there are any $\omega, \omega' \in \{0,1\}^V$ such that $\omega(i) = \omega'(i)$ for all $i \in I$ but $S_\omega \neq S_{\omega'}$, then for any $x_V^0 \cup x_V^1$, $\chi_{S_\omega}(x_V^\omega) = \chi_{S_\omega}(x_I^\omega) = \chi_{S_\omega}(x_I^{\omega'}) \neq \chi_{S_{\omega'}}(x_I^{\omega'}) = \chi_{S_{\omega'}}(x_V^{\omega'})$. In particular, one of these two values must be 0, so the whole product is 0.

So we may restrict to the case where S_ω depends only on $\omega \upharpoonright I$. Let v be the unique element in $V \setminus I$ and let $E' = E \upharpoonright \binom{I}{k}$. Then we have the decomposition

$$\int \cdot d\mu_E^{V+V} = \iint \cdot d\mu_{E, x_I^0 \cup x_I^1}^{v+v} d\mu_{E'}^{I+I} = \iiint \cdot d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_{E'}^{I+I}.$$

The second equality holds because the graph E^{V+V} used to defined the measure μ_E^{V+V} does not contain any edges containing both $(v, 0)$ and $(v, 1)$. So we have

$$\begin{aligned} & \int \prod_{\omega \in \{0,1\}^V} (f\chi_{S_\omega})(x_V^\omega) d\mu_E^{V+V} \\ &= \iint \prod_{\omega \in \{0,1\}^V} (f\chi_{S_\omega})(x_V^\omega) d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_{E'}^{I+I} \\ &= \iint \prod_{\omega \in \{0,1\}^I} f\chi_{S_\omega}(x_I^\omega, x_v^0) \prod_{\omega \in \{0,1\}^I} f\chi_{S_\omega}(x_I^\omega, x_v^1) d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_{E'}^{I+I} \\ &= \int \left(\int \prod_{\omega \in \{0,1\}^I} f\chi_{S_\omega}(x_I^\omega, x_v) d\mu_{E, x_I^0 \cup x_I^1}^v \right)^2 d\mu_{E'}^{I+I} \end{aligned}$$

Since the inside of the integral is always non-negative, this term is non-negative. \square

In particular, since $\int \prod_{\omega \in \{0,1\}^V} f(x_V^\omega) d\mu_E^{V+V} \geq 0$, $\|f\|_{U_\infty^V(\mu_E^V)}$ is defined.

Next we want a Cauchy-Schwarz style inequality for these seminorms:

Lemma 6.4 (Gowers-Cauchy-Schwarz). *Suppose that for each $\omega \in \{0,1\}^V$, f_ω is an $L^\infty(\mu_E^V)$ function. Then*

$$\left| \int \prod_{\omega \in \{0,1\}^V} f_\omega(x_V^\omega) d\mu_E^{V+V} \right| \leq \prod_{\omega \in \{0,1\}^V} \|f_\omega\|_{U_\infty^V(\mu_E^V)}.$$

Proof. Fix some $v \in V$, and let $I = V \setminus \{v\}$. Note that we have the decomposition

$$\int \cdot d\mu_E^{V+V} = \iint \cdot d\mu_{E, x_I^0 \cup x_I^1}^{v+v} d\mu_E^{I+I} = \iiint \cdot d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_{E, x_I^0 \cup x_I^1}^v d\mu_E^{I+I}.$$

As above, the second equality holds because the graph in μ_E^{V+V} does not contain any edges containing both $(v, 0)$ and $(v, 1)$. For $\omega \in \{0, 1\}^I$ and $b \in \{0, 1\}$, let us write ωb for the element of $\{0, 1\}^V$ given by $(\omega b)(i) = \omega(i)$ if $i \in I$ and $(\omega b)(i) = b$ if $i = v$. Therefore, using Cauchy-Schwarz, we have:

$$\begin{aligned} & \left| \int \prod_{\omega \in \{0, 1\}^V} f_\omega(x_V^\omega) d\mu_E^{V+V} \right|^2 \\ &= \left| \int \left(\int \prod_{\omega \in \{0, 1\}^V} f_\omega(x_I^\omega, x_v^{\omega(v)}) d\mu_{E, x_I^0 \cup x_I^1}^{v+v} \right) d\mu_E^{I+I} \right|^2 \\ &= \left| \int \left(\int \prod_{\omega \in \{0, 1\}^I} f_{\omega 0}(x_I^\omega, x_v^0) d\mu_{E, x_I^0 \cup x_I^1}^v \right) \left(\int \prod_{\omega \in \{0, 1\}^I} f_{\omega 1}(x_I^\omega, x_v^1) d\mu_{E, x_I^0 \cup x_I^1}^v \right) d\mu_E^{I+I} \right|^2 \\ &\leq \int \left(\int \prod_{\omega \in \{0, 1\}^I} f_{\omega 0}(x_I^\omega, x_v) d\mu_{E, x_I^0 \cup x_I^1}^v \right)^2 d\mu_E^{I+I} \int \left(\int \prod_{\omega \in \{0, 1\}^I} f_{\omega 1}(x_I^\omega, x_v) d\mu_{E, x_I^0 \cup x_I^1}^v \right)^2 d\mu_E^{I+I} \\ &\leq \int \prod_{\omega \in \{0, 1\}^V} f_{(\omega \upharpoonright I)0}(x_V^\omega) d\mu_E^{V+V} \int \prod_{\omega \in \{0, 1\}^V} f_{(\omega \upharpoonright I)1}(x_V^\omega) d\mu_E^{V+V} \end{aligned}$$

In particular, applying this repeatedly to each coordinate in V , we have

$$\begin{aligned} \left| \int \prod_{\omega \in \{0, 1\}^V} f_\omega(x_V^\omega) d\mu_E^{V+V} \right|^{2^V} &\leq \prod_{\omega \in \{0, 1\}^V} \int \prod_{\omega' \in \{0, 1\}^V} f_\omega(x_V^{\omega'}) d\mu_E^{V+V} \\ &= \prod_{\omega \in \{0, 1\}^V} \|f_\omega\|_{U_\infty^V(\mu_E^V)}. \end{aligned}$$

□

Corollary 6.5. $|\int f d\mu_E^V| \leq \|f\|_{U_\infty^V(\mu_E^V)}.$

Proof. In the previous lemma, take $f_{0^V} = f$ and $f_\omega = 1$ for $\omega \neq 0^V$. □

Lemma 6.6. $\|\cdot\|_{U_\infty^V(\mu_E^V)}$ is a seminorm.

Proof. It is easy to see from the definition that $\|cf\|_{U_\infty^V(\mu_E^V)} = |c| \cdot \|f\|_{U_\infty^V(\mu_E^V)}.$

To see subadditivity, observe that $\|f+g\|_{U_\infty^V(\mu_E^V)}^{2^{|V|}}$ expands to a sum of $2^{2^{|V|}}$ integrals, each of which, by the previous lemma, is bounded by $\|f\|_{U_\infty^V(\mu_E^V)}^m \|g\|_{U_\infty^V(\mu_E^V)}^{2^{|V|}-m}$

for a suitable m . In particular, this bound is precisely $\left(\|f\|_{U_\infty^V(\mu_E^V)} + \|g\|_{U_\infty^V(\mu_E^V)}\right)^{2^{|V|}}$ as desired. \square

In particular, the work above gives:

Theorem 6.7. *If $\|\mathbb{E}(f \mid \mathcal{B}_{V,<V})\|_{L^2(\mu_E^V)} > 0$ then $\|f\|_{U_\infty^V(\mu_E^V)} > 0$.*

Proof. If $\|\mathbb{E}(f \mid \mathcal{B}_{V,<V})\|_{L^2(\mu_E^V)} > 0$ then we may find, for each $I \subseteq V$ with $|I| = |V| - 1$, $B_I \in \mathcal{B}_{V,I}$ such that

$$0 < \left| \int f \prod_I \chi_{B_I} d\mu_E^V \right| \leq \|f \prod_I \chi_{B_I}\|_{U_\infty^V(\mu_E^V)}.$$

By repeatedly applying Lemma 6.3, once to each I , we have

$$0 < \|f \prod_I \chi_{B_I}\|_{U_\infty^V(\mu_E^V)} \leq \|f\|_{U_\infty^V(\mu_E^V)}.$$

\square

We will obtain the converse, which will show that $\|\mathbb{E}(f \mid \mathcal{B}_{V,<V})\|_{L^2(\mu_E^V)} > 0$ iff $\|f\|_{U_\infty^V(\mu_E^V)} > 0$, and in particular will enable us to show that μ has J -regularity.

6.2. Characterization in Product Measures.

Definition 6.8. We say μ_E^V is a *product measure* if no element of E contains more than one element of V .

(Recall that μ_E^V abbreviates μ_{E,a_P}^V , so there may be edges in E connecting elements of V to elements of P .) We call such measures product measures because they are extensions of the ordinary product measure $\prod_{v \in V} \mu_E^v$.

We are not concerned with the converse in the case where $P \neq \emptyset$, so we state it only when $E \subseteq \binom{V}{k}$.

Theorem 6.9. *If μ_E^V is a product measure, and $\|\mathbb{E}(f \mid \mathcal{B}_{V,<V})\|_{L^2(\mu_E^V)} = 0$ then $\|f\|_{U_\infty^V(\mu_E^V)} = 0$.*

Proof. This is essentially identical to the argument we gave for regularity for ordinary measures. Suppose $\|\mathbb{E}(f \mid \mathcal{B}_{V,<V})\|_{L^2(\mu^V)} = 0$. We have

$$\begin{aligned} \|f\|_{U_\infty^V(\mu_E^V)} &= \int f(x_V^0) \prod_{\omega \in \{0,1\}^V, \omega \neq 0^V} f(x_V^\omega) d\mu_E^{V+V} \\ &= \iint f(x_V^0) \prod_{\omega \in \{0,1\}^n, \omega \neq 0^V} f(x_V^\omega) d\mu_E^V d\mu_E^V \end{aligned}$$

This last equality holds because μ_E^V is a product measure, and so the inner copy of μ_E^V does not depend on the choice of x_V^1 .

Observe that, for every particular value of x_V^1 , $\prod_{\omega \in \{0,1\}^V, \omega \neq 0^V} f(x_V^\omega)$ is $\mathcal{B}_{V, < V}$ -measurable, and therefore

$$\int f(x_V^0) \prod_{\omega \in \{0,1\}^V, \omega \neq 0^V} f(x_V^\omega) d\mu_E^V = 0.$$

□

6.3. Seminorms for Non-Principal Algebras. We will need a more general family of seminorms corresponding to arbitrary algebras of the form $\mathcal{B}_{V, \mathcal{I}}$.

Definition 6.10. For $J \subseteq V$, define

$$\|f\|_{U_\infty^{V,J}(\mu_E^V)} = \left(\int \prod_{\omega \in \{0,1\}^J} f(x_{V \setminus J}, x_J^\omega) d\mu_E^{V+J} \right)^{2^{-|J|}}.$$

Note that $\|f\|_{U_\infty^{V,V}(\mu_E^V)} = \|f\|_{U_\infty^V(\mu_E^V)}$.

We need to generalize to norms $U_\infty^{V, \mathcal{J}}$ where \mathcal{J} is a set. A natural choice would be to take the product of $U_\infty^{V,J}$ over all $J \in \mathcal{J}$, but this is not a seminorm. Instead we need the following form:

Definition 6.11. Let $\mathcal{J} \subseteq \mathcal{P}(V)$ be a set such that if $J, J' \in \mathcal{J}$ are distinct then $J \not\subseteq J'$. Then we define

$$\|f\|_{U_\infty^{V, \mathcal{J}}(\mu_E^V)} = \inf \sum_{i \leq k} \left(\prod_{J \in \mathcal{J}} \|f_i\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} \right)^{\frac{1}{\sum_{J \in \mathcal{J}} 2^{|J|}}}$$

where the infimum is taken over all sequences f_0, \dots, f_k such that $f = \sum_{i \leq k} f_i$.

It is not immediately obvious that $U_\infty^{V,J}$ and $U_\infty^{V, \{J\}}$ calculate the same value, but this will follow immediately once we show that $U_\infty^{V,J}$ is a seminorm.

Lemma 6.12. *If f is an $L^\infty(\mu_E^V)$ function then*

$$0 \leq \int \prod_{\omega \in \{0,1\}^J} f(x_{V \setminus J}, x_J^\omega) d\mu_E^{V+J}.$$

Proof. Let $V' = V \setminus J$. We have

$$\begin{aligned} \int \prod_{\omega \in \{0,1\}^J} f(x_{V'}, x_J^\omega) d\mu_E^{V+J} &= \iint \prod_{\omega \in \{0,1\}^J} f(x_{V'}, x_J^\omega) d\mu_{E, x_{V'}}^{J+J} d\mu_E^{V'} \\ &= \int \|f(x_{V'}, \cdot)\|_{U_\infty^J(\mu_{E, x_{V'}}^J)}^{2^{|J|}} d\mu_E^{V'} \\ &\geq 0. \end{aligned}$$

□

Lemma 6.13. $|\int f d\mu_E^V| \leq \|f\|_{U_\infty^{V,\mathcal{J}}(\mu_E^V)}$

Proof. First consider the case where \mathcal{J} is a singleton $\{J\}$. Again, let $V' = V \setminus J$.

$$\begin{aligned} \left| \int f d\mu_E^V \right|^{2^{|J|}} &= \left| \iint f d\mu_{E,x_{V'}}^J d\mu_E^{V'} \right|^{2^{|J|}} \\ &\leq \int \left| \int f d\mu_{E,x_{V'}}^J \right|^{2^{|J|}} d\mu_E^{V'} \\ &\leq \int \|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^J} d\mu_E^{V'} \\ &= \|f\|_{U_\infty^{V,\mathcal{J}}(\mu_E^V)}^{2^{|J|}}. \end{aligned}$$

For the general case, first observe that, setting $c = \sum_{J \in \mathcal{J}} 2^{|J|}$,

$$\begin{aligned} \left| \int f d\mu_E^V \right|^c &= \prod_{J \in \mathcal{J}} \left| \int f d\mu_E^V \right|^{2^{|J|}} \\ &\leq \prod_{J \in \mathcal{J}} \|f\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}}. \end{aligned}$$

So if $f = \sum_{i \leq k} f_i$ we have

$$\left| \int f d\mu_E^V \right| \leq \sum_{i \leq k} \left| \int f_i d\mu_E^V \right| \leq \sum_{i \leq k} \left(\prod_{J \in \mathcal{J}} \|f_i\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} \right)^{\frac{1}{c}}.$$

This holds for any $\sum_{i \leq k} f_i$, so $|\int f d\mu_E^V| \leq \|f\|_{U_\infty^{V,\mathcal{J}}(\mu_E^V)}$. \square

Lemma 6.14. $\|\cdot\|_{U_\infty^{V,\mathcal{J}}(\mu_E^V)}$ is a seminorm.

Proof. Once again positive homogeneity is obvious from the definition, so we need only check that the triangle inequality holds.

We first consider the case where \mathcal{J} is a singleton:

$$\begin{aligned} \|f + g\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} &= \int \|f + g\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}} d\mu_E^{V'} \\ &\leq \int (\|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}} + \|g\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}}) d\mu_E^{V'} \\ &= \|f\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} + \|g\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} \\ &\leq (\|f\|_{U_\infty^{V,J}(\mu_E^V)} + \|g\|_{U_\infty^{V,J}(\mu_E^V)})^{2^{|J|}}. \end{aligned}$$

For $|\mathcal{J}| > 1$, we may use the fact that if $f = \sum_{i \leq k} f_i$ and $g = \sum_{j \leq m} g_j$ then $f + g = \sum_{i \leq k} f_i + \sum_{j \leq m} g_j$. \square

The main thing that makes the uniformity seminorms useful to us is that they easily pass across different measures:

Lemma 6.15. *Let $J \subseteq V$ and $V' = V \setminus J$. If $\|f\|_{U_\infty^J(\mu_E^J)} = 0$ then for $\mu_E^{V'}$ -almost-every $x_{V'}$, $\|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)} = 0$.*

Proof.

$$\begin{aligned}
0 &= \|f\|_{U_\infty^J(\mu_E^J)}^{2^{|J|}} \\
&= \int \prod_{\omega \in \{0,1\}^J} f(x_J^\omega) d\mu_E^J \\
&= \int \prod_{\omega \in \{0,1\}^J} f(x_J^\omega) \int 1 d\mu_{E,x_J^0 \cup x_J^1}^{V'} d\mu_E^{J+J} \\
&= \int \prod_{\omega \in \{0,1\}^J} f(x_J^\omega) d\mu_{E,x_V}^{J+J} d\mu_E^{V'} \\
&= \int \|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}} d\mu_E^{V'}.
\end{aligned}$$

Therefore for $\mu_E^{V'}$ -almost-every $x_{V'}$, $\|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)} = 0$. \square

In order to associate these more general seminorms with the correct algebras, we introduce the following definition:

Definition 6.16. If $\mathcal{I} \subseteq \mathcal{P}(V)$ is non-empty, we define \mathcal{I}^\perp to be the set of $J \subseteq V$ such that:

- (1) There is no $I \in \mathcal{I}$ with $J \subseteq I$,
- (2) If $J' \subsetneq J$ then there is an $I \in \mathcal{I}$ with $J' \subseteq I$.

$$J^- = \{I \subseteq V \mid J \not\subseteq I\}.$$

\cdot^\perp and \cdot^- depend on the choice of the ambient set V . Note that \mathcal{I}^\perp always has the property that if $J, J' \in \mathcal{I}^\perp$ are distinct then $J \not\subseteq J'$, so $\|\cdot\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)}$ is defined. Also, if $\mathcal{I} = \{I \subseteq V \mid |I| = |V| - 1\}$ then $\mathcal{I}^\perp = \{V\}$. Also, note that $(J^-)^\perp = \{J\}$.

We will show that when μ_E^V is nice enough, $\mathcal{B}_{V,\mathcal{I}}$ and $\bigcap_{J \in \mathcal{I}^\perp} \mathcal{B}_{V,J^-}$ agree up to μ_E^V measure 0.

Lemma 6.17. *If there is no $J \in \mathcal{J}$ such that $J \subseteq I$ and B is \mathcal{B}_{V,I^-} -measurable then*

$$\|f\chi_B\|_{U_\infty^{V,\mathcal{J}}(\mu_E^V)} \leq \|f\|_{U_\infty^{V,\mathcal{J}}(\mu_E^V)}.$$

Proof. It suffices to show this for \mathcal{J} a singleton $\{J\}$. Write $V' = V \setminus J$. Observe that for any fixed $x_{V'}$, the function $\chi_B(x_{V'}, \cdot)$ is a $\mathcal{B}_{J,J \cap I^-}$ -measurable

function, where $J \cap I$ must be a proper subset of J . So we have:

$$\begin{aligned} \|f\chi_B\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} &= \int \|f\chi_B\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}} d\mu_E^{V'} \\ &\leq \int \|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}} d\mu_E^{V'} \\ &= \|f\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}}. \end{aligned}$$

□

Theorem 6.18. *If $\|\mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}})\|_{L^2(\mu_E^V)} > 0$ then $\|f\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)} > 0$.*

Proof. If $\|\mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}})\|_{L^2(\mu_E^V)} > 0$ then we may find, for each $I \in \mathcal{I}$, a set $B_I \in \mathcal{B}_{V,I}$, such that

$$0 < \left| \int f \prod_I \chi_{B_I} d\mu_E^V \right| \leq \|f \prod_I \chi_{B_I}\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)}.$$

Observe that for each $I \in \mathcal{I}$ we may apply the previous lemma, so we have

$$0 < \|f \prod_I \chi_{B_I}\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)} \leq \|f\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)}.$$

□

Definition 6.19. Let μ be a canonical family of measures of size k and degree $\sum_{I \in E} 2^{|I \cap V|}$. For some $\mathcal{I} \subseteq \mathcal{P}(V)$, we say $U_E^{V,\mathcal{I}^\perp}(\mu_E^V)$ is *characteristic* if for each f , $\|f\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)} > 0$ implies $\|\mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}})\|_{L^2(\mu_E^V)} > 0$.

Lemma 6.20. *If $U_\infty^{J,\mathcal{I}}(\mu_E^J)$ is characteristic for each $\mathcal{I} \subseteq \mathcal{P}(J)$ then μ_E^V has J -regularity.*

Proof. Let $J \subsetneq V$ and $\mathcal{I} \subseteq \mathcal{P}(V)$ be given, and let g and f_I be as in the definition of regularity. Let $h = g - \mathbb{E}(g \mid \mathcal{B}_{V,\mathcal{I} \wedge J}) = g - \mathbb{E}(g \mid \mathcal{B}_{J,\mathcal{I} \wedge J})$ (viewing g as a function on \mathcal{B}_J). Since $\|\mathbb{E}(h \mid \mathcal{B}_{J,\mathcal{I} \wedge J})\|_{L^2(\mu_E^J)} = 0$, by assumption we have $\|h\|_{U_\infty^{J,\mathcal{I} \wedge J}(\mu_E^J)} = 0$. Then we also have $\|h\|_{U_\infty^{J,\mathcal{I} \wedge J}(\mu_{E,x_{V \setminus J}}^J)} = 0$ for $\mu_E^{V \setminus J}$ -almost-every $x_{V \setminus J}$. (The exact choice of *which* set of measure 1 this holds on depends on the choice of representative of h .)¹

¹We note the similarity of this argument to the one in [21]. The argument there uses two equivalent characterizations of a regularity type property, DISC and PAIR, the former analogous to having 0 projection and the latter to having 0 uniformity norm; a key step in that DISC implies PAIR in the dense setting, PAIR in the dense setting implies PAIR in the sparse setting, and then PAIR in the sparse setting implies DISC.

Including $x_{V \setminus J}$ as part of the background parameters, Theorem 6.18 implies that $\|\mathbb{E}(h \mid \mathcal{B}_{J, \mathcal{I} \wedge J})\|_{L^2(\mu_{E, x_{V \setminus J}}^J)} = 0$, and so

$$\begin{aligned} \int h \prod_I f_I d\mu_E^V &= \int h \prod_I f_I d\mu_{E, x_{V \setminus J}}^J d\mu_E^{V \setminus J} \\ &= 0 \end{aligned}$$

since for almost every fixed $x_{V \setminus J}$, $\prod_I f_I$ is $\mathcal{B}_{J, \mathcal{I} \wedge J}$ -measurable. \square

6.4. Characterization for Non-Principal Seminorms. The principal seminorms are the controlling case for showing that the uniformity norms are characteristic: in this subsection we show that if the principal algebra of a given size is characterized by its uniformity norm then all algebras of the same size are characterized by their uniformity norms. We only need this for the case of a product measure, but we include the general argument for completeness.

Lemma 6.21. *Let \mathcal{I} be given and let $J \subseteq V$. If μ_E^V has J -regularity then $\mathcal{B}_{V, \mathcal{I} \wedge \{J\}} = \mathcal{B}_{V, \mathcal{I}} \cap \mathcal{B}_{V, J}$ up to μ_E^V -measure 0.*

Proof. By definition, we have $\mathcal{B}_{V, \mathcal{I} \wedge \{J\}} \subseteq \mathcal{B}_{V, \mathcal{I}} \cap \mathcal{B}_{V, J}$.

For the converse, we first show that if g is $\mathcal{B}_{V, J}$ -measurable with $\mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I} \wedge \{J\}}) = 0$, also $\mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I}}) = 0$. Let such a g be given, and for each $I \in \mathcal{I}$, let f_I be $\mathcal{B}_{V, I}$ -measurable, so $\prod_I f_I$ is $\mathcal{B}_{V, \mathcal{I}}$ -measurable. Since $g = g - \mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I} \wedge \{J\}})$ and μ_E^V has J -regularity,

$$\int g \prod_I f_I d\mu_E^V = 0,$$

showing that $\mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I}}) = 0$.

Now let g $\mathcal{B}_{V, \mathcal{I}} \cap \mathcal{B}_{V, J}$ -measurable be given. It suffices to show that $g' = g - \mathbb{E}(g \mid \mathcal{B}_{V, \mathcal{I} \wedge \{J\}}) = 0$. But g' is $\mathcal{B}_{V, \mathcal{I}} \cap \mathcal{B}_{V, J}$ -measurable and satisfies $\mathbb{E}(g' \mid \mathcal{B}_{V, \mathcal{I} \wedge \{J\}}) = 0$, and therefore satisfies $\mathbb{E}(g' \mid \mathcal{B}_{V, \mathcal{I}}) = g' = 0$. \square

Lemma 6.22. *For any $\mathcal{I}, \mathcal{J} \subseteq \mathcal{P}(V)$, if μ_E^V has J -regularity for every $J \in \mathcal{J}$ then $\mathcal{B}_{V, \mathcal{I} \wedge \mathcal{J}}$ is $\mathcal{B}_{V, \mathcal{I}} \cap \mathcal{B}_{V, \mathcal{J}}$ up to μ_E^V -measure 0.*

Proof. The direction $\mathcal{B}_{V, \mathcal{I} \wedge \mathcal{J}} \subseteq \mathcal{B}_{V, \mathcal{I}} \cap \mathcal{B}_{V, \mathcal{J}}$ is immediate from the definition.

For the converse, we may assume $\mathcal{J} = \{J_1, \dots, J_n\}$ where $i \neq j$ implies $J_i \not\subseteq J_j$, and we proceed by induction on n . When $n = 1$ this is just the previous lemma. Suppose the claim holds for \mathcal{J} and we wish to show it for $\mathcal{J} \cup \{J\}$. Note that

$$\mathcal{B}_{\mathcal{I} \wedge (\mathcal{J} \cup \{J\})} = \mathcal{B}_{(\mathcal{I} \wedge \mathcal{J}) \cup (\mathcal{I} \wedge \{J\})} = \mathcal{B}_{\mathcal{I} \wedge \mathcal{J}} \uplus \mathcal{B}_{\mathcal{I} \wedge \{J\}}.$$

It suffices to show that whenever f is $\mathcal{B}_{V,\mathcal{I}}$ -measurable then $\mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I} \cup \{J\}})$ is $\mathcal{B}_{V,\mathcal{I} \wedge (\mathcal{I} \cup \{J\})}$ -measurable. For any f , we have

$$\begin{aligned} \mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I} \cup \{J\}}) &= \mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}} \uplus \mathcal{B}_J) \\ &= \mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}}) + \mathbb{E}(f \mid \mathcal{B}_J) - \mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}} \cap \mathcal{B}_J). \end{aligned}$$

When f is $\mathcal{B}_{V,\mathcal{I}}$ -measurable, $\mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}}) - \mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}} \cap \mathcal{B}_J)$ is $\mathcal{B}_{V,\mathcal{I}} \cap \mathcal{B}_{V,\mathcal{I}}$ -measurable, and therefore, by IH, $\mathcal{B}_{V,\mathcal{I} \wedge \mathcal{I}}$ -measurable. By the previous lemma, $\mathbb{E}(f \mid \mathcal{B}_J)$ is $\mathcal{B}_{V,\mathcal{I} \wedge \{J\}}$ -measurable. In particular, this means $\mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I} \cup \{J\}})$ is $\mathcal{B}_{\mathcal{I} \wedge \mathcal{I}} \uplus \mathcal{B}_{\mathcal{I} \wedge \{J\}} = \mathcal{B}_{\mathcal{I} \wedge (\mathcal{I} \cup \{J\})}$ -measurable. \square

Lemma 6.23. *If μ_E^V has J -regularity for every $J \in \mathcal{I}^\perp$, $\mathcal{B}_{V,\mathcal{I}}$ is $\bigcap_{J \in \mathcal{I}^\perp} \mathcal{B}_{V,J}$ -up to μ_E^V -measure 0.*

Proof. We have $\bigcap_{J \in \mathcal{I}^\perp} \mathcal{B}_{V,J}$ is $\mathcal{B}_{V,\bigwedge_{J \in \mathcal{I}^\perp} J}$ -up to μ_E^V -measure 0 (it is easy to see that \wedge is associative and commutative, so this follows by repeated application of Lemma 6.22). We need only check that $\bigwedge_{J \in \mathcal{I}^\perp} J^- = \mathcal{I}$.

If $I \in \mathcal{I}$ (or even $I \subseteq I' \in \mathcal{I}$) then for every $J \in \mathcal{I}^\perp$, we have $J \not\subseteq I$, and therefore $I \in J^-$, and therefore $I \in \bigwedge_{J \in \mathcal{I}^\perp} J^-$. Conversely, if there is no $I' \in \mathcal{I}$ such that $I \subseteq I'$ then there is a $J \subseteq I$ such that $J \in \mathcal{I}^\perp$, and therefore no $J' \in J^-$ such that $I \subseteq J'$, and therefore $I \notin \bigwedge_{J \in \mathcal{I}^\perp} J^-$. \square

Note that the following theorem is one of the places where we directly appeal to the definability structure of our σ -algebras. This is for a good reason: the statement would not be true if we replaced our σ -algebras with, say, simple product algebras.

Theorem 6.24. *Suppose that for every $J \in \mathcal{I}^\perp$ and every $J' \subseteq J$, $U_\infty^{J'}(\mu_E^{J'})$ is characteristic and that for $\mu_E^{V \setminus J'}$ -almost-every $x_{V'}$, $U_\infty^{J'}(\mu_{E,x_{V'}}^{J'})$ is characteristic. Then $U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)$ is characteristic.*

Proof. We proceed by main induction on $|V|$. In particular, if $V \in \mathcal{I}^\perp$ then the claim is given by the assumption, so we may assume that every element $J \in \mathcal{I}^\perp$ has $|J| < |V|$, and so by IH, each $U_\infty^{J,\mathcal{I}^\perp}(\mu_E^J)$ is characteristic.

We start with the case where $\mathcal{I}^\perp = \{J\}$. Suppose $\|f\|_{U_\infty^{V,J}(\mu_E^V)} > 0$, so, setting $V' = V \setminus J$, also

$$0 < \|f\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} = \int \|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}} d\mu_{E'}^{V'}.$$

Since almost every $U_\infty^J(\mu_{E,x_{V'}}^J)$ is characteristic, there must be a set $S_0 \subseteq M^{V'}$ of positive measure such that, for $x_{V'} \in S_0$, $\|f\|_{U_\infty^J(\mu_{E,x_{V'}}^J)}^{2^{|J|}} > 0$. By assumption, for almost every $x_{V'} \in S_0$, we have $\|\mathbb{E}(f \mid \mathcal{B}_{J,<J})\|_{L^2(\mu_{E,x_{V'}}^J)} > 0$. This means that for almost every $x_{V'} \in S_0$, we may choose a set $B(x_{V'}, b_{Q^{x_{V'}}}^{x_{V'}}) \in \mathcal{B}_{J,<J}^0$ such that $|\int f \chi_{B(x_{V'})} d\mu_{E,x_{V'}}^J| > 0$. (Note that here

even the index set $Q^{x_{V'}}$ may depend on $x_{V'}$.) There are only countably many formulas, so we may assume that there is a single formula defining $B(x_{V'}, b_Q^{x_{V'}})$, independently of $x_{V'} \in S_0$. However there are uncountably many choices of parameters, so we may *not* assume that the parameters $b_Q^{x_{V'}}$ are independent of $x_{V'}$. We may choose an $\epsilon > 0$, an approximation of f by a simple function f' , and a set $S_1 \subseteq S_0$ of positive measure so that for $x_{V'} \in S_1$, $|\int f' \chi_{B(x_{V'}, b_Q^{x_{V'}})} d\mu_{E, x_{V'}}^J| \geq \epsilon$.

Recall the distinguished function symbols $\mathbf{max}_Q^{E, J, f', B}$; these symbols choose values of the parameters in ψ maximizing the value of $|\int f' \chi_{B(x_{V'}, b_Q^{x_{V'}})} d\mu_{E, x_{V'}}^J|$.

So, replacing $B(x_{V'}, b_Q^{x_{V'}})$ with $\hat{B}(x_{V'}, a_W) = B(x_{V'}, \mathbf{max}_Q^{E, J, f', B}(x_{V'}, a_W))$ (where a_W are the parameters in the definition of f'),

$$\left| \int f' \chi_{\hat{B}(x_{V'}, a_W)} d\mu_{E, x_{V'}}^J \right| \geq \left| \int f' \chi_{B(x_{V'}, b_Q^{x_{V'}})} d\mu_{E, x_{V'}}^J \right|.$$

In particular, for each $x_{V'} \in S_1$, $\left| \int f' \chi_{\hat{B}(x_{V'}, a_W)} d\mu_{E, x_{V'}}^J \right| \geq \epsilon$. Note that $\hat{B}(x_{V'}, a_W) \in \mathcal{B}_{J, < J}^0$ and therefore $\hat{B}(a_W) \in \mathcal{B}_{J, J-}^0$.

We may partition $S_1 = S_1^+ \cup S_1^-$ where $x_{V'} \in S_1^+$ exactly when $\int f' \chi_{\hat{B}(x_{V'}, a_W)} d\mu_{E, x_{V'}}^J \geq \epsilon$. Clearly at least one of S_1^+ and S_1^- has measure $\geq \mu_E^{V'}(S_1)/2$; without loss of generality, we assume S_1^+ does. Since f' is simple, we have $f' = \sum_{i \leq n} \alpha_i \chi_{C_i}$. We may write a large union D of sets consisting of those $x_{V'}$ such that

$$\begin{aligned} & \left(m_{E, x_{V'}}^J(C_1(x_{V'}, a_W) \cap \hat{B}(x_{V'}, a_W)) < \beta_1 \wedge m_{E, x_{V'}}^J(C_1(x_{V'}, a_W) \cap \hat{B}(x_{V'}, a_W)) > \beta'_1 \right) \\ & \wedge \dots \\ & \wedge \left(m_{E, x_{V'}}^J(C_n(x_{V'}, a_W) \cap \hat{B}(x_{V'}, a_W)) < \beta_n \wedge m_{E, x_{V'}}^J(C_n(x_{V'}, a_W) \cap \hat{B}(x_{V'}, a_W)) > \beta'_n \right) \end{aligned}$$

so that $\mu_E^{V'}(D \cap S_1^+) \geq (1 - \delta)\mu_E^{V'}(S_1^+)$ and every element of D satisfies

$$\int f' \chi_{\hat{B}(x_{V'}, a_W)} d\mu_{E, x_{V'}}^J > \epsilon/2.$$

The formula defining this set has only free variables $a_W, x_{V'}$, so D is $\mathcal{B}_{V, V'}$ -measurable. Then

$$\int f' \chi_{\hat{B}(x_{V'}, a_W)} \chi_D d\mu_E^V d\mu = \iint f' \chi_{\hat{B}(x_{V'}, a_W)} \chi_D d\mu_{E, x_{V'}}^J d\mu_E^{V'} > \epsilon(1 - \delta)\mu_E^{V'}(S_1)/2.$$

Since we chose f' to be an arbitrarily close approximation of f , we may assume that $\|f - f'\|_{L^2(\mu_E^V)} < \epsilon(1 - \delta)\mu_E^{V'}(S_1)/4$, and so we have

$$\int f \chi_{\hat{B}(x_{V'}, a_W)} \chi_D d\mu_E^V d\mu > \epsilon(1 - \delta)\mu_E^{V'}(S_1)/4 > 0.$$

Since $\chi_{\hat{B}(x_{V'}, a_W)} \chi_D$ is $\mathcal{B}_{V, J-}$ -measurable, we are finished.

For the general case, suppose $\|\mathbb{E}(f \mid \mathcal{B}_{V,\mathcal{I}})\|_{L^2(\mu_E^V)} = 0$. By Lemma 6.20, μ_E^V has J -regularity for each $J \in \mathcal{I}^\perp$, so by Lemma 6.23, $\mathcal{B}_{V,\mathcal{I}} = \bigcap_{J \in \mathcal{I}^\perp} \mathcal{B}_{V,J^-}$, and so $\|\mathbb{E}(f \mid \bigcap_{J \in \mathcal{I}^\perp} \mathcal{B}_{V,J^-})\|_{L^2(\mu_E^V)} = 0$. Let $\mathcal{I}^\perp = \{J_1, \dots, J_r\}$. Then we may define a sequence $f_0 = f$, $f_{i+1} = \mathbb{E}(f_i \mid \mathcal{B}_{V,J_{i+1}^-})$,

$$f = f_r + (f_{r-1} - f_r) + (f_{r-2} - f_{r-1}) + \dots + (f_0 - f_1).$$

Since $f_r = \mathbb{E}(f \mid \bigcap_{J \in \mathcal{I}^\perp} \mathcal{B}_{V,J^-})$, we have $f_r = 0$. We therefore have

$$\|f\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)} \leq \sum_{i < r} \left(\prod_{J \in \mathcal{I}^\perp} \|f_i - f_{i+1}\|_{U_\infty^{V,J}(\mu_E^V)}^{2^{|J|}} \right)^{\frac{1}{\sum_{J \in \mathcal{I}^\perp} 2^{|J|}}}.$$

For each $i < r$, $f_i - f_{i+1} = f_i - \mathbb{E}(f_i \mid \mathcal{B}_{V,J_{i+1}^-})$. In particular, $\|\mathbb{E}(f_i - f_{i+1} \mid \mathcal{B}_{V,J_{i+1}^-})\|_{L^2(\mu_E^V)} = 0$, and therefore by the previous part, $\|f_i - f_{i+1}\|_{U_\infty^{V,J_{i+1}}(\mu_E^V)} = 0$. But this means the whole sum is 0, and therefore $\|f\|_{U_\infty^{V,\mathcal{I}^\perp}(\mu_E^V)} = 0$. \square

7. RANDOM MEASURES

Theorem 7.1. *Suppose μ is a canonical family of measures of size k and degree $\sum_{I \in E} 2^{2|I|}$. Then $U_\infty^V(\mu_E^V)$ is characteristic.*

Proof. We show a more general result:

Let $I \subseteq V$ be given and let $V' = V \setminus I$. Let

$$F(x_{V'}, x_I^0, x_I^1) = \prod_{\omega \in \{0,1\}^I} f_\omega(x_{V'}, x_I^\omega)$$

be such that

$$0 < \int F(x_{V'}, x_I^0, x_I^1) d\mu_E^{V+I}.$$

Then $\|\mathbb{E}(f_{0^I} \mid \mathcal{B}_{V,<V})\|_{L^2(\mu_E^V)} > 0$.

The main result is then the case where $I = V$ and $f_{\emptyset,\omega} = f$ for all ω .

We proceed by induction on $|I|$. When $I = \emptyset$, this is trivial, so assume $|I| > 0$.

Fix some $v \in I$, and let $I' = I \setminus \{v\}$. For each $\omega \in \{0,1\}^{I'}$ and each $b \in \{0,1\}$ we will write ωb for the corresponding elements of $\{0,1\}^I$. We define a function

$$G(x_{V'}, x_{I'}^0, x_{I'}^1) = \int \prod_{\omega \in \{0,1\}^{I'}} f_{\omega 1}(x_{V'}, x_{I'}^\omega, x_v^1) d\mu_{E, x_{V'} \cup x_{I'}^0, x_{I'}^1}^v.$$

Let W be the vertex set of the measure $\mu_E^{V \setminus \{v\} + I'}$; that is, $W = V' \cup (I' \times \{0,1\})$. Let E' be the corresponding hypergraph, so $\mu_E^{V \setminus \{v\} + I'} = \mu_{E'}^W$. Each edge I of E corresponds to at most $2^{|I \cap V|}$ edges of E' , so $\sum_{I \in E'} 2^{|I' \cap W|} \leq \sum_{I \in E} 2^{2|I|}$. (There is likely some room here for optimizing the exact degree

of the canonical family needed.) Let $\mathcal{J} \subseteq \mathcal{P}(W)$ be the collection of subsets of the form

$$V' \cup \{(i, \omega(i)) \mid i \in I'\}$$

for some $\omega \in \{0, 1\}^{I'}$. That is, \mathcal{J} consists of those sets which contain V' together with exactly one copy of each coordinate from I' . The elements of \mathcal{J}^\perp are pairs $J = \{(i, 0), (i, 1)\}$ for some $i \in I'$. No edge of E' contains both elements of one of the pairs $\{(i, 0), (i, 1)\}$, so $\mu_{E'}^J$ and $\mu_{E', x_{W \setminus J'}}^J$ are product measures, and in particular, $U_\infty^J(\mu_{E'}^J)$ and $U_\infty^J(\mu_{E', x_{W \setminus J'}}^J)$ are characteristic by Theorem 6.9.

We claim that G is $\mathcal{B}_{W, \mathcal{J}}$ -measurable (with respect to the measure $\mu_{E'}^W$). Suppose H is a function with $\|\mathbb{E}(H \mid \mathcal{B}_{W, \mathcal{J}})\|_{L^2(\mu_{E'}^W)} = 0$. By Theorem 6.24, $U_\infty^{W, \mathcal{J}^\perp}(\mu_{E'}^W)$ is characteristic, so $\|H\|_{U_\infty^{W, \mathcal{J}^\perp}(\mu_{E'}^W)} = 0$, and therefore for $\mu_{E'}^v$ -almost-every x_v^1 , $\|H\|_{U_\infty^{W, \mathcal{J}^\perp}(\mu_{E', x_v^1}^W)} = 0$, and so $\|\mathbb{E}(H \mid \mathcal{B}_{W, \mathcal{J}})\|_{L^2(\mu_{E', x_v^1}^W)} = 0$. Then

$$\begin{aligned} & \int H(x_{V'}, x_{I'}^0, x_{I'}^1) \cdot G(x_{V'}, x_{I'}^0, x_{I'}^1) d\mu_{E'}^W \\ &= \iint H \prod_{\omega \in \{0, 1\}^{I'}} f_{\omega 1}(x_{V'}, x_{I'}^\omega, x_v^1) d\mu_{E', x_{V'} \cup x_{I'}^0 \cup x_{I'}^1}^v d\mu_{E'}^W \\ &= \iint H \prod_{\omega \in \{0, 1\}^{I'}} f_{\omega 1}(x_{V'}, x_{I'}^\omega, x_v^1) d\mu_{E', x_v^1}^W d\mu_E^v \\ &= 0 \end{aligned}$$

Since this holds for any H with $\|\mathbb{E}(H \mid \mathcal{B}_{W, \mathcal{J}})\|_{L^2(\mu_{E'}^W)} = 0$, it follows that G is $\mathcal{B}_{W, \mathcal{J}}$ -measurable. This means that we may write

$$G(x_{V'}, x_{I'}^0, x_{I'}^1) = \lim_{N \rightarrow \infty} \sum_{n \leq N} \prod_{\omega \in \{0, 1\}^{I'}} g_{\omega, n, N}(x_{V'}, x_{I'}^\omega)$$

in the $L^2(\mu_{E'}^W)$ -norm. We may assume the $g_{\omega, n, N}$ are $L^\infty(\mu_{E'}^W)$ functions.

Then we have some ϵ such that

$$\begin{aligned} 0 < \epsilon &< \int \prod_{\omega \in \{0, 1\}^{I'}} f_\omega(x_{V'}, x_{I'}^\omega) d\mu_E^{V+I} \\ &= \int \prod_{\omega \in \{0, 1\}^{I'}} f_{\omega 0}(x_{V'}, x_{I'}^\omega, x_v^0) \prod_{\omega \in \{0, 1\}^{I'}} f_{\omega 1}(x_{V'}, x_{I'}^\omega, x_v^1) d\mu_E^{V+I} \\ &= \int \prod_{\omega \in \{0, 1\}^{I'}} f_{\omega 0}(x_{V'}, x_{I'}^\omega, x_v^0) G(x_{V'}, x_{I'}^0, x_{I'}^1) d\mu_E^{V+I}. \end{aligned}$$

Choosing N large enough, we may make

$$\begin{aligned} & \|G(x_{V'}, x_{I'}^0, x_{I'}^1) - \sum_{n \leq N} \prod_{\omega \in \{0,1\}^{I'}} g_{\omega,n,N}(x_{V'}, x_{I'}^\omega)\|_{L^2(\mu_{E'}^W)} \\ & < \frac{\epsilon}{2 \prod_{\omega \in \{0,1\}^{I'}} \|f_{\omega 0}(x_{V'}, x_I^\omega, x_v^0)\|_{L^\infty(\mu_E^{V+I'})}} \end{aligned}$$

and therefore

$$\begin{aligned} 0 < \epsilon/2 & < \int \prod_{\omega \in \{0,1\}^{I'}} f_{\omega 0}(x_{V'}, x_I^\omega, x_v^0) \sum_{n \leq N} \prod_{\omega \in \{0,1\}^{I'}} g_{\omega,n,N}(x_{V'}, x_{I'}^\omega) d\mu_E^{V+I'} \\ & = \sum_{n, N} \int \prod_{\omega \in \{0,1\}^{I'}} f_{\omega 0}(x_{V'}, x_I^\omega, x_v^0) g_{\omega,n,N}(x_{V'}, x_{I'}^\omega) d\mu_E^{V+I'} \end{aligned}$$

In particular, there must be some n such that

$$0 < \int \prod_{\omega \in \{0,1\}^{I'}} f_{\omega 0}(x_{V'}, x_I^\omega, x_v^0) g_{\omega,n,N}(x_{V'}, x_{I'}^\omega) d\mu_E^{V+I'}.$$

Consider the functions given by, for each $\omega \in \{0,1\}^{I'}$, setting $f'_\omega = f_{\omega 0} g_{\omega,n,N}$. We apply IH to I' , and conclude that $\|\mathbb{E}(f'_{0I'} \mid \mathcal{B}_{V,<V})\|_{L^2(\mu_E^V)} > 0$. Since $g_{0I',n,N}$ is $\mathcal{B}_{V,V' \cup I'} \subseteq \mathcal{B}_{V,<V}$ -measurable, it follows that $\|\mathbb{E}(f_{0I'} \mid \mathcal{B}_{V,<V})\|_{L^2(\mu_E^V)} > 0$ as well. \square

We can now give a sparse version of the hypergraph removal lemma:

Theorem 7.2. *For every k -uniform hypergraph K on vertices V and every constant $\epsilon > 0$, there are δ, ζ so that whenever Γ is a $\zeta, |K|2^{2k}$ -suitably random k -uniform hypergraph and $A_n \subseteq \Gamma_n$ with $\frac{\text{hom}(K, A_n)}{|\Gamma_n^K|} < \delta$ then there is a subset L of A with $|L| \leq \epsilon |\Gamma|$ such that $\text{hom}(K, A \setminus L) = 0$.*

Proof. Suppose not. Let K, ϵ be a counterexample. Since there are no such δ, ζ , for each n we may choose k -uniform hypergraphs $H_n \subseteq \Gamma_n$ with Γ_n $1/n, |K|2^{2k}$ -suitably random and $\frac{\text{hom}(K, H_n)}{|\Gamma_n^K|} < 1/n$. Let \mathfrak{M} be the model given by Theorem 5.5.

Let V be the set of vertices of K . For any $J \subseteq V$, Theorem 7.1 implies that $U_\infty^J(\mu_K^J)$ is characteristic, and therefore by Lemma 6.20, μ_E^V has J -regularity. For each $I \in K$, let $A_I = \{x_V \mid x_I \in A\}$. Since $\frac{\text{hom}(K, A_n)}{|\Gamma_n^K|} \rightarrow 0$, we have $\mu_K^V(\bigcap_{I \in K} A_I) = 0$. Then by Theorem 4.2, there must be definable sets B_I such that $\mu_K^I(A_I \setminus B_I) < \epsilon/|K|$ and $\bigcap_{I \in K} B_I = \emptyset$. Let $C = \bigcup_{I \in K} (A_I \setminus B_I)$, so $\mu_{\{1,k\}}^{[1,k]}(C) < \epsilon$. C is definable from parameters in M , and therefore the properties $\mu_{\{1,k\}}^{[1,k]}(C) < \epsilon$ and $\bigcap_{I \in K} (A_I \setminus C) = \emptyset$ are witnessed by formulas. Therefore there must be arbitrarily large finite models where these formulas are satisfied. But this contradicts the choice of the hypergraphs H_n, Γ_n . \square

8. CONCLUSION

The notion of suitable randomness used in this paper is strong compared to other notions of pseudorandomness for hypergraphs that have been considered [3, 4, 6, 22]. The next step towards developing a rich analytic approach to working with sparse random hypergraphs would be a detailed investigation of the relationship between notions of pseudorandomness in the finite setting and the corresponding properties of measures in the infinitary setting. With less than suitable randomness, we would expect to lose the full Fubini property, but the notions that replace it are likely to be of interest themselves.

The approach Conlon and Gowers use to prove hypergraph regularity [5] depends, like our approach, on the use of various norms to detect the presence of certain properties. Their norms are much more narrowly tailored than the general uniformity norms. The uniformity norms are strikingly natural in the infinitary setting, lining up with canonical algebras of definable sets; it is possible that other norms also correspond to algebras which might be of independent interest.

REFERENCES

- [1] T. Austin and T. Tao. Testability and repair of hereditary hypergraph properties. *Random Structures & Algorithms*, pages 376–463, 2010. 10.1002/rsa.20300.
- [2] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztegombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.
- [3] F. R. K. Chung. Quasi-random classes of hypergraphs. *Random Structures & Algorithms*, 1(4):363–382, 1990.
- [4] F. R. K. Chung and R. L. Graham. Quasi-random hypergraphs. *Random Structures & Algorithms*, 1(1):105–124, 1990.
- [5] D. Conlon and W. Gowers. Combinatorial theorems in sparse random sets. <http://www.dpmms.cam.ac.uk/~dc340/SparseRandom.pdf>, 2011.
- [6] D. Conlon, H. Hàn, Y. Person, and M. Schacht. Weak quasi-randomness for uniform hypergraphs. <http://www.dpmms.cam.ac.uk/~dc340/wqr.pdf>, 2009.
- [7] G. Elek and B. Szegedy. Limits of hypergraphs, removal and regularity lemmas. a non-standard approach. unpublished, 2007.
- [8] P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs Combin.*, 2(2):113–121, 1986.
- [9] J. Fox. A new proof of the graph removal lemma. *Ann. of Math. (2)*, 174(1):561–579, 2011.
- [10] P. Frankl and V. Rödl. Extremal problems on set systems. *Random Structures & Algorithms*, 20(2):131–164, 2002.
- [11] E. Friedgut, V. Rödl, and M. Schacht. Ramsey properties of random discrete structures. *Random Structures & Algorithms*, 37(4):407–436, 2010.
- [12] R. Goldblatt. *Lectures on the hyperreals*, volume 188 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998. An introduction to nonstandard analysis.
- [13] I. Goldbring and H. Towsner. An approximate logic for measures. *ArXiv e-prints*, June 2011.
- [14] W. T. Gowers. A new proof of Szemerédi’s theorem. *Geometric And Functional Analysis*, 11(3):465–588, Aug 2001. 10.1007/s00039-001-0332-9.

- [15] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math. (2)*, 166(3):897–946, 2007.
- [16] E. Hrushovski. Stable group theory and approximate subgroups. *J. Amer. Math. Soc.*, 25(1):189–243, 2012.
- [17] H. J. Keisler. Hyperfinite model theory. In *Logic Colloquium 76 (Oxford, 1976)*, pages 5–110. Studies in Logic and Found. Math., Vol. 87. North-Holland, Amsterdam, 1977.
- [18] H. J. Keisler. An infinitesimal approach to stochastic analysis. *Mem. Amer. Math. Soc.*, 48(297):x+184, 1984.
- [19] Y. Kohayakawa. Szemerédi’s regularity lemma for sparse graphs. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pages 216–230. Springer, Berlin, 1997.
- [20] Y. Kohayakawa, T. Łuczak, and V. Rödl. Arithmetic progressions of length three in subsets of a random set. *Acta Arith.*, 75(2):133–163, 1996.
- [21] Y. Kohayakawa, V. Rödl, M. Schacht, and J. Skokan. *An Irregular Mind (Szemerédi is 70)*, volume 21, chapter On the triangle removal lemma for subgraphs of sparse pseudorandom graphs, pages 359–404. Bolyai Soc. Math. Stud., 2010.
- [22] Y. Kohayakawa, V. Rödl, and J. Skokan. Hypergraphs, quasi-randomness, and conditions for regularity. *J. Combin. Theory Ser. A*, 97(2):307–352, 2002.
- [23] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.
- [24] L. Lovász and B. Szegedy. Szemerédi’s lemma for the analyst. *Geom. Funct. Anal.*, 17(1):252–270, 2007.
- [25] B. Nagle, V. Rödl, and M. Schacht. The counting lemma for regular k -uniform hypergraphs. *Random Structures & Algorithms*, 28(2):113–179, 2006.
- [26] V. Rödl and J. Skokan. Regularity lemma for k -uniform hypergraphs. *Random Structures & Algorithms*, 25(1):1–42, 2004.
- [27] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, Vol. II, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 939–945. North-Holland, Amsterdam, 1978.
- [28] M. Schacht. Extremal results for discrete random structure. http://www.math.uni-hamburg.de/home/schacht/preprints/extremal_rs.pdf, 2009.
- [29] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.
- [30] T. Tao. A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A*, 113(7):1257–1280, 2006.
- [31] T. Tao. A correspondence principle between (hyper)graph theory and probability theory, and the (hyper)graph removal lemma. *J. Analyse Math.*, 103:1–45, 2007. 10.1007/s11854-008-0001-0.
- [32] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems*, 28(2):657–688, 2008.
- [33] H. Towsner. Convergence of diagonal ergodic averages. *Ergodic Theory Dynam. Systems*, 29(4):1309–1326, 2009. 10.1017/S0143385708000722.
- [34] H. Towsner. A model theoretic proof of Szemerédi’s theorem. *ArXiv e-prints*, Feb. 2010.
- [35] H. Towsner. A correspondence principle for the Gowers uniformity norms. *Logic and Analysis*, 4, February 2012.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CONNECTICUT U-3009, 196 AUDI-
TORIUM ROAD, STORRS, CT 06269-3009, USA

E-mail address: `henry.towsner@uconn.edu`

URL: `www.math.uconn.edu/~towsner`